A MODELING FRAMEWORK FOR ANALYZING THE EDUCATION SYSTEM AS A COMPLEX SYSTEM

A Thesis
Presented to
The Academic Faculty

by

Pratik Mital

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
H. Milton Stewart School of Industrial and Systems Engineering

Georgia Institute of Technology
May 2015

A MODELING FRAMEWORK FOR ANALYZING THE EDUCATION SYSTEM AS A COMPLEX SYSTEM

Approved by:

Professor Donna C. Llewellyn, Advisor

Center for the Enhancement of Teaching and Learning
Georgia Institute of Technology (Retired)

Institute for STEM and Diversity Initiatives
Boise State University

Professor David M. Goldsman, Co-Advisor

H. Milton Steward School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Joel Sokol

H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Professor Christiaan J. Paredis

George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology

Professor Gordon A. Kingsley

School of Public Policy
Georgia Institute of Technology

Date Approved: March 9, 2015

To my grandparents Sushma Gupta, Surendra Kumar Gupta, Sadhana Mital, and Rajendra Kumar Mital

And

To my parents Anju Mital and Pankaj Mital

ACKNOWLEDGEMENTS

This thesis would not have been possible without the guidance, support, and contribution of the following people. I sincerely convey my thanks to each one of them.

First, and foremost, my advisor Dr. Donna Llewellyn. She has always been by my side, and her support has been invaluable. She encouraged me to always probe one step further, provided constructive criticism, and created an environment where I could freely share my ideas and thoughts about research, job search, and anything else that was on my mind. She sincerely cared about my progress during the thesis process and always kept the big picture in mind. This helped me plan the thesis in a very productive way. Her enthusiasm towards this work kept me going even during some of the roadblocks that we faced during this process. She always took out time to give me feedback, and I cannot imagine a better advisor than her.

My co-advisor Dr. David Goldsman. His support towards this thesis and his inputs were key factors in the completion of this thesis. He stepped up to be my co-advisor when Dr. Llewellyn was retiring. He has always had my best interest in mind and I am grateful for that. His jovial nature and enthusiasm, no matter what the situation, is something that I admire a lot.

My thesis committee members: Dr. Gordon Kingsley, Dr. Chris Paredis, and Dr. Joel Sokol. This is an inter-disciplinary thesis and each of their backgrounds truly represents that. Coming from public policy, systems engineering, and industrial engineering, the committee members provided their own unique perspective towards this work. The numerous discussions that we have had and their inputs have improved this thesis a lot.

My colleague, Dr. Roxanne Moore, with whom I have worked closely since the beginning of this thesis. We have written numerous papers together and I hope we continue doing so. Her inputs towards this thesis have been critical in its completion and she has always been sincerely interested in improving the quality of this thesis. I have thoroughly enjoyed working with her.

The educational research team from CEISMC at Georgia Tech, specially, Dr. Marion Usselman, Mike Ryan, and Douglas Edwards. This thesis would not have been complete without the inputs and data that I have received from CEISMC. The constant process of having back and forth discussions with the educational research team helped in building a thesis that can be used in practice.

During the first two years of my Ph.D. program, I was supported by Dr. Marc Goetschalckx as a research and teaching assistant. I am grateful for this support. These were my first steps as a Ph.D. student, and this experience helped in shaping me as a more applied researcher in industrial engineering.

Mrs. Pamela Morrison, for taking care of all the administrative work that a dissertation process involves. She was the first person I met in the department when I first came to Georgia Tech, and she has always been sincerely interested in my progress throughout.

My undergraduate professor Dr. Kiran Seth, for introducing me towards operations research. He always encouraged me to pursue higher studies, and nudged me on this exciting and challenging path of research.

My parents, grandparents, and other family members who have always encouraged me, believed in me, and supported me. It is because of their nurturing in my formative years that I have reached this stage in my life. They always gave me the freedom to do what I truly like and always stood by me in any decision that I took.

My better half, Aditi Sharma. I met her 9 years ago, and she has been an integral part of my life since then. She has always kept me on the right path, and made me a better person, both professionally and otherwise. Her sincerity and honesty is what I admire the most. Her support has been invaluable throughout the dissertation process.

My best friends Harshit and Viru. I am truly blessed to have such great friends. They always have time for me to discuss my problems and to share both the happy and sad moments of my life. Their belief in me gave me confidence to keep going. Their numerous visits to Atlanta made my stay here truly memorable.

My badminton family, Ankit, Varun, and many others who are amongst my most cherished friends and with whom I have shared so many important moments of my life. My undergraduate friends, Sudhanshu, and Mayank, who always had my best interest at heart and have encouraged and supported me since the day we met.

My mentor from the ISyE department at Georgia Tech, Aly, for his guidance throughout the Ph.D. process. His advice has been invaluable to me, and I really appreciate him being sincerely interested in my progress. There are very few people who put in so much time and effort in others, and I was fortunate to have him on my side.

My ex-roommate, Vishal, who introduced me to object oriented programming and helped me improve my programming skills. This has played an integral part in implementing my research. Another ex-roommate, Nabankur. He has been a great friend and has always supported me. He truly cherishes my happiness and has been there for me for the past five years.

My friends from the ISyE department at Georgia Tech, Isil, Bahar, and Diego. They made my time in the department really enjoyable. We have shared so many fond memories together. Also, Dibyajat, Dhruv, and many others from Georgia Tech, with whom I have shared many special moments. It would not be the same without them.

Finally, the grant which funded this research, National Science Foundation (NSF) award 1238089, Advanced Manufacturing and Prototyping Integrated to Unlock Potential (AMP-IT-UP).

Table of Contents

ACKNOWLEDGEMENTS iv

LIST OF TABLES x

LIST OF FIGURES xi

SUMMARY xv

CHAPTER 1: INTRODUCTION...1

1.1 Is the Education System Complex? 3

1.2 Outline..6

CHAPTER 2: LITERATURE REVIEW 8

2.1 Systems Engineering Approaches for Modeling the Education System 8

2.2 Social Network Analysis in the Education System...................................10

2.3 Interventions in the School System..10

2.4 Agent-Based Modeling Frameworks...13

CHAPTER 3: EWB CASE STUDY FOR FRAMEWORK DEVELOPMENT.........15

3.1 Introduction to EWB...15

3.2 Model to Analyze the EWB Case Study..16

3.3 Results and Analyses 28

3.4 Validation..38

CHAPTER 4: MODELING FRAMEWORK: ESIM..40

4.1 Introduction to ESIM 40

4.2 Model Definition..44

4.3 Model Design 65

4.4 Model Analysis 80

4.5 Model Validation 84

CHAPTER 5: APPLICATION OF ESIM TO SLIDER 89

5.1 Introduction to SLIDER 89

5.2 Application of ESIM 91

5.2.1 Model Definition 91

5.2.2 Model Design 99

5.2.3 Model Analysis 107

5.2.4 Model Validation 128

5.3 Modeling at a Higher Scale 131

CHAPTER 6: CONCLUSIONS AND FUTURE WORK 135

6.1 EWB Case Study 135

6.2 Overview of ESIM 137

6.3 SLIDER Case Study 138

6.4 Key Insights 140

6.5 Policy Implications 141

6.6 Future Avenues of Research 144

APPENDIX A: SIMULATION RESULTS: SLIDER 149

REFERENCES 159

LIST OF TABLES

1 Complex system characteristics 4

2 Agents' behaviors for the EWB case study 22

3 Mean and standard deviation of model outputs 29

4 Sensitivity analysis with respect to weights 32

5 Sensitivity analysis with respect to the number of time periods 34

6 Sensitivity analysis with respect to the S-curve parameter 35

7 Effect of scale on modeling change in the system 48

8 Agents and their attributes 50

9 Teacher attribute assessment rubric 52

10 Performance of $P(I)$ for various gap levels 64

11 Agents' behaviors 66

12 Agents' behaviors for the SLILDER case study 100

13 Mean and variance in the final fitness output 108

14 Probability of sustainability of the intervention at School 1 and 2 121

15 Sensitivity analysis with respect to internal and external weights 122

16 Sensitivity analysis with respect to the transient phase parameter 124

LIST OF FIGURES

1 Agent network for EWB case study 19

2 Method of Morris, total money generated 37

3 Inter-disciplinary collaboration for framework development 41

4 Policy cycle as a consequence of applying the framework 42

5 Modeling cycle 43

6 Hierarchical structure of the education system 45

7 Modeling at the scale of a school 46

8(a) Acceptable zone - Realized 57

8(b) Acceptable zone – Not realized 57

9 3-dimensional acceptable zone using the gaps 58

10 Graph of $a(x)$ 62

11 Attribute state transition between consecutive time steps 75

12 Relationship state transition between consecutive time steps 77

13 Agent network for SLIDER case study 93

14(a) School 1 8th grade CRCT scores 96

14(b) School 2 8th grade CRCT scores 96

15(a) School 1 student-teacher ratio 97

15(b) School 2 student-teacher ratio 98

16(a) Change in Teacher 1's support for intervention 109

16(b) Change in Teacher 1's inquiry teaching skill 110

16(c) Change in Teacher 1's content knowledge 110

16(d) Change in Teacher 1's self-efficacy 111

17(a) Change in Class 1's SLIDER test scores in year 3 111

17(b) Change in Class 1's SLIDER test scores in year 4 112

18 Change in Administration's support for intervention 112

19(a) Policy management gap in School 1 113

19(b) Capability gap in School 1 113

19(c) Cultural gap in School 1 114

19(d) Overall fitness of School 1 114

20(a) Change in Teacher 2's support for intervention 115

20(b) Change in Teacher 2's inquiry teaching skill 116

20(c) Change in Teacher 2's content knowledge 116

20(d) Change in Teacher 2's self-efficacy 117

21(a) Change in Class 2's SLIDER test scores in year 3 117

21(b) Change in Class 2's SLIDER test scores in year 4 118

22 Change in Administration's support for intervention 118

23(a) Policy management gap in School 2 119

23(b) Capability gap in School 2 119

23(c) Cultural gap in School 2 120

23(d) Overall fitness of School 2 120

24(a) Barriers & enablers for School 1 125

24(b) Barriers & enablers for School 2 126

25 MoM – Probability of sustainability 128

26 SLIDER network at a macro-level 131

27 Proportion of additional support towards the three schools 133

28(a) Change in Teacher 2's support for intervention 149

28(b) Change in Teacher 2's inquiry teaching skill 149

28(c) Change in Teacher 2's content knowledge 150

28(d) Change in Teacher 2's self-efficacy 151

29(a) Change in Class 2's SLIDER test scores in year 3 151

29(b) Change in Class 2’s SLIDER test scores in year 4 152

30(a) Change in Teacher 1’s support for intervention....................................... 152

30(b) Change in Teacher 1’s inquiry teaching skill... 153

30(c) Change in Teacher 1’s content knowledge 153

30(d) Change in Teacher 1’s self-efficacy... 154

31(a) Change in Class 1’s SLIDER test scores in year 3 154

31(b) Change in Class 1’s SLIDER test scores in year 4 155

32(a) Change in Teacher 3’s support for intervention....................................... 155

32(b) Change in Teacher 3’s inquiry teaching skill... 156

32(c) Change in Teacher 3’s content knowledge 156

32(d) Change in Teacher 3’s self-efficacy... 157

33(a) Change in Class 3’s SLIDER test scores in year 3 157

33(b) Change in Class 3’s SLIDER test scores in year 4 158

SUMMARY

In this thesis, we introduce the Education System Intervention Modeling Framework (ESIM Framework), for analyzing interventions in the K-12 education system. This framework is the first of its kind to model interventions in the K-12 school system in the United States.

Techniques from systems engineering and operations research, such as agent-based modeling and social network analysis, are used to model the bottom-up mechanisms of intervention implementation in schools. This is a new application of this methodology in a domain which is of considerable importance at the local, state, and federal levels where large dollar amounts are being invested. By applying the ESIM framework, an intervention can be better analyzed in terms of the barriers and enablers to intervention implementation and sustainability. The risk of failure of future interventions is thereby reduced through improved allocation of resources towards the system agents and attributes which play key roles in the sustainability of the intervention. Increasing the sustainability of interventions in the school system improves educational outcomes in the schools and increases the benefits gained from the millions of dollars being invested in such interventions.

The ESIM framework presented in this thesis contains detailed steps for modeling an intervention in a school setting. Developing a framework has many advantages: it can be applied across different types of interventions, it facilitates repeatability of results and comparisons across the models, and it provides a detailed step-by-step method of documenting the model-building process. Application of the framework also requires

collaboration from public policy researchers, educational researchers, and practitioners of the interventions, which are captured in the framework methodology.

The framework was developed with the help of a case study of an extracurricular school intervention, an *Engineers Without Borders* chapter, implemented in a magnet school setting through a partnership with the Georgia Institute of Technology, Atlanta (Georgia Tech) as part of a National Science Foundation (NSF) GK-12 grant. This case study is ideal for the development of the framework and as its first application because it had two different outcomes over two different years, which helped in developing insights about the success of this intervention. Also, the scale of this intervention was small enough to test the development and application of the framework. With the help of this case study, a more-generalized framework is developed which is applicable across a broad range of education system interventions.

The ESIM framework developed is divided into four phases: model definition, model design, model analysis, and model validation. In the <u>model definition</u> phase, the overview of the problem to be modeled is documented. Then, detailed descriptions about the agents, attributes, and the environment being modeled are provided. Other modeling decisions, such as scale and time horizons, are also made in this phase. Finally, the criteria for a sustainable intervention is defined along with a method to quantify the risk of implementing the intervention in the particular school system. In the <u>model design</u> phase, the conceptual model is built using agent-based modeling, social network analysis, and discrete-time Markov chains. Then, the conceptual model is validated with the help of subject matter experts (SMEs) using Pace's 4C's framework for conceptual model validation. After that, the computer simulation model is implemented and verified. In the model analysis phase, simulation results are generated and analyzed. While simulating outcomes that are consistent with reality is helpful, the real contributions of this framework are two-fold: the sensitivity analysis of the model, and the determination of factors that are likely to affect the intervention outcomes. The latter is accomplished using the Method of Morris, a factorial sampling technique. Finally, in the model validation phase, verification and validation techniques are applied. This step is critical in developing confidence in the model amongst its users.

The ESIM framework is then applied to a case study of a curriculum intervention, *Science Learning: Integrating Design, Engineering and Robotics,* involving the design and implementation of an 8th-grade, inquiry-based physical science curriculum across three demographically varying schools. This intervention was also implemented in collaboration with Georgia Tech as part of an NSF DRK-12 grant. This was a five year intervention from Sep 2009 to Oct 2014. This case study provides a good comparison of the implementation of the intervention across different school settings because of the varied outcomes at the three schools.

As a future application, the ESIM framework has been embedded in the NSF-funded EarSketch Math-Science Partnership, a computer science intervention that will be implemented across 30 different schools, reaching about 50 teachers and more than 1000 students over the 4 -year grant period.

CHAPTER 1

INTRODUCTION

"All models are wrong, but some are useful" - George E.P. Box

Educational interventions and reforms are commonplace, but only a limited number prove to be truly effective. From controversial federal policies such as 'No Child Left Behind' [1], to individual teachers adopting new pedagogical techniques, it can be difficult to measure the success of educational interventions and even more difficult to understand why they fail. The complexity of a school or school system is easily underestimated from the top-down view; curricula and interventions are often designed outside of the intended school settings or are copied from other schools, only to fail in context. There are many factors that can impact the outcome of an intervention, including the cognitive abilities of individual students, the grit of a teacher, the socioeconomic status of the community, the principal's pedagogical beliefs, the standardized testing regime, the school schedule, the level of parental involvement, and so on. Hence, we understand that schools and school districts are complex, dynamic systems affected by numerous factors. However, most educational policy makers and reformers have yet to rely on quantitative models when making resource allocation decisions for school interventions. While no single model is applicable across a broad spectrum of school contexts, a unified *framework* can be applied to build models for particular settings. Ultimately, by applying the framework to diverse school settings, models are developed and analyzed to identify common attributes and relationships that are likely to help or hinder intervention implementation [2][1].

In this thesis, a modeling *framework* has been developed that will help policy analysts identify and understand barriers and enablers for educational interventions in different school settings. This framework has been named the Education System Intervention Modeling Framework (ESIM). This is a new application of industrial and systems engineering in a not-for-profit domain where millions of dollars are invested at the local, state, and federal levels. The education system is a complex system and demonstrating that industrial and systems engineering can improve the effectiveness of interventions in this system is quite valuable. Through the application of ESIM, the risk of failure of an intervention can be reduced by improved allocation of resources towards those components of the system which play key roles in the success of that intervention. Increasing the success of interventions in the school system improves the educational outcomes in the school and increases the benefits gained from the investments being made in such interventions.

The major research objectives of this thesis are the following:

1. Study techniques from industrial engineering, systems engineering and operations research which can be used to analyze interventions in the education system.
2. Develop a framework which is applicable across a broad range of interventions and can be used to develop models for them.

[1] This paragraph has been adapted from our work in [2]

3. Utilize existing conceptual frameworks from educational and public policy research into the framework being developed in this thesis.
4. Quantify sustainability/risk of implementing an intervention in a given school system.
5. Model different case studies and gain insights into their successful implementation.

These research objectives are addressed in different chapters of this thesis. Chapter 2 looks at prior research conducted in modeling the education system. Chapter 3 uses a case study to develop the framework and presents the model developed for that case study. Chapter 4 presents the ESIM framework, which utilizes different techniques from industrial engineering, systems engineering and operations research. It also utilizes conceptual frameworks from educational and public policy research. A probability measure is developed as part of the framework to quantify sustainability/risk of implementing an intervention. Chapters 5 presents another case study to which the ESIM framework is applied.

In the next section, the complexity of the education system is discussed. Even though the education system clearly looks like a complex system, it is useful to explicitly define the various components in the education system which make it complex. This helps in providing insights about how to develop a framework for analyzing the behavior of the system.

1.1 Is the Education System Complex?

Complex system science is a field that has gained a great deal of momentum in the past decade and has been used to analyze engineered systems, healthcare, economics, military conflicts and ethnic violence [3]. The complex systems discipline has emerged out of the application of industrial and systems engineering, operations research, and network theory to analyze systems which exhibit complexity. A system can be either simple or complex depending upon its characteristics. The following table from Sterman [4] discusses the properties which make a system complex.

Table 1: Complex systems characteristics

Characteristics of complex systems
• **Constantly changing**: Change in the system occurs at many time scales and levels of representation. • **Tightly coupled**: The actors in the system interact strongly with one another and with the environment. • **Governed by feedback**: Because of the tight couplings among actors, actions feedback on themselves and alter the system state. Dynamics arise from these feedback loops. • **Nonlinear**: Nonlinearity arises as multiple factors interact in decision making. • **History-dependent**: Many actions are irreversible and future system states depend upon where you are right now. However, the system does not have to be history-path dependent necessarily, and it may be reasonable to assume that the current state of the system only depends upon the previous state of this system. • **Self-organizing**: The dynamics of systems arise spontaneously from their internal structure. • **Adaptive**: The capabilities and decision rules of the agents in complex systems change over time.

Table 1 continued

• **Delays in feedback**: Time delays in feedback channels mean the long-run response of a system to an intervention is often different from its short-run response. • **Emergent**: Properties at the micro-level lead to emergent properties of the system at the macro-level. • **Policy resistant**: The complexity of the systems in which we are embedded overwhelms our ability to understand them. The result: Many seemingly obvious solutions to problems fail or actually worsen the situation.

The education system clearly exhibits all the above properties of a complex system. It is a dynamic system whose agents and the environment interact with each other and change over time. The relationships between the students, teachers, principal, community, and administrative and government bodies play a major role on the performance of the system and change the system state over time. The interaction amongst the agents in the educational system results in system-level dynamics which are hard to understand by looking at the parts of the system in isolation. There are emergent patterns observed in the system as a consequence of the system dynamics taking place, where emergent behavior is defined as the large-scale behavior observed as a result of the interactions taking place at the smaller scale. The concept of scale is introduced while developing the modeling framework in Chapter 3. Any intervention or action taken by the agents affects the other agents and this creates feedback loops in the system. The feedback being received, however, has delays, and this can cause an intervention which seems to improve the educational system in the short run to eventually cause more harm than good in the long run. Another consequence of feedback is that the agents are constantly adapting themselves with respect to changes in the environment. Because there are many confounding factors that affect the system performance, the effect that the agents have on each other is hard to quantify and is highly nonlinear. The system also has memory; for example, a student whose test scores are low cannot generally move to high scores in the next time period. All of these factors make the education system a dynamic and integrated multilayer system of people, money, knowledge, and information – in other words, a 'complex system'. Hence, rigorous techniques from industrial engineering, systems engineering, and operations research are required to analyze and understand the properties of the education system.

1.2 Outline

In this thesis, the ESIM framework is presented for creating models of various school system interventions and analyzing the critical factors that cause the particular intervention to succeed or fail. These models aid in the improvement of the education system by providing quantitative insights about the important attributes, relationships, and resources affecting the success of the interventions. The framework makes use of techniques such as agent-based modeling and social network analysis, as well as other tools from industrial and systems engineering to model the system.

The remainder of the thesis is organized as follows: In Chapter 2, a literature review is presented which establishes the state of the art in applying techniques from industrial and systems engineering, as well as operations research to model the education system. A brief review of the application of social network analysis techniques in the education system is presented after that. Educational and policy literature is also surveyed related to analyzing interventions in the school system. Some agent-based modeling frameworks that have been developed for ecology, civil engineering, and social systems are also reviewed and a discussion is provided about how those can be helpful in developing a framework for the education system. In Chapter 3, the case study of an *Engineers Without Borders* (EWB) intervention, which was used in the development of the modeling framework, is presented. The model developed, to analyze the barriers and enablers for a successful implementation of this intervention, is also presented in this chapter. In Chapter 4, the ESIM framework is presented. The framework includes detailed steps to create, analyze, verify, and validate a model for a specific school system intervention. The framework has four phases: model definition, model design, model analysis, and model validation, all of which are discussed in detail in their respective sections. In Chapter 5, the ESIM framework is then applied to a case study of the *Science Learning: Integrating, Design, Engineering and Robotics* (SLIDER) intervention implemented across three schools. Simulation results and analyses are presented, which enable deeper understanding of the drivers for intervention success and sustainability in these interventions. Finally in Chapter 6, conclusions and future work are discussed, including the insights gained from applying the framework to the case studies and possible future extensions of this thesis. Appendix at the end contains some additional simulation results for the SLIDER case study presented in Chapter 5.

CHAPTER 2

LITERATURE REVIEW

In the previous chapter, the need for the ESIM framework was discussed along with the characteristics of the education system which make it complex. Major research objectives of the thesis were also presented and an outline of the thesis was provided.

In this chapter, the focus is on the relevant literature surrounding the techniques used in this thesis to analyze the education system. Prior work in this area is surveyed, along with a critical review. Conceptual studies about educational interventions are also discussed along with some of the agent-based modeling frameworks developed for other fields of application.

2.1 Systems Engineering Approaches for Modeling the Education System

Educational researchers have long studied school reform and the issues of what facilitates and hinders success in curricular and other interventions [5, 6]. Experts in educational policy and public policy also have studied the interaction of policies and practices of reform agendas within social and organizational contexts [7-9]. However, these approaches either look only at specific parts of the system or lack the analytical tools to analyze the dynamics taking place at the systems level. Industrial and systems engineering, which had its origins in studying manufacturing systems, is a field where researchers have made great contributions towards understanding complex systems, including transportation systems [10], supply chain networks [11], financial systems [12], health care [13], humanitarian support systems [14], and even the education system as discussed below. The work cited for each discipline is just a representative pub and there are many such examples. However, there is a need to apply more-rigorous systems engineering and operations research techniques to model the education system [15].

Recent advances in this area include the application of system dynamics (SD) and agent-based modeling (ABM) to simulate US student interest and selection in STEM (Science, Technology, Engineering and Mathematics) [16, 17]. These models are a good starting point, but they take a top-down view of the education system and do not adequately capture grass-roots mechanisms at the school level. Groff takes a dynamic complex systems perspective to analyze the education system and promotes the use of system dynamics as an analysis tool [18]. While some system dynamics-based school-level models exist to investigate policy impacts on enrollment and academic performance of students, these models rely heavily on survey data to formulate causal relationships without a mechanism for distinguishing correlation from causation [19, 20].

Agent-based modeling on the other hand, has emerged as a popular technique to model and simulate human systems, and the following three statements capture its advantages over other techniques: *(i)* It captures emergent phenomena; *(ii)* It provides a natural description of a system; and *(iii)* It is flexible [21]. It allows for human entities in the system and their characteristics to be modeled as agents and their attributes respectively. The relationships amongst the agents define the agent-based modeling environment. Through these relationships, exchange of resources can be simulated along with the change in the attributes of the interacting agents.

2.2 Social Network Analysis in the Education System

Beyond system dynamics and agent-based modeling, social network analysis can be an important tool for understanding a particular school environment. Educational researchers have realized the importance of social networks in the education system and have begun to analyze the effects of the social network structure in implementing interventions [22]. Different studies have analyzed teacher networks and their effect on the implementation of instructional reforms [23, 24]. Daly [25] also uses different case studies of educational interventions to illustrate the changes in teacher networks over time and the impact of teacher networks on the implementation of an intervention. Thus far, however, social network analysis of school systems in the education literature has been done in isolation from the modeling techniques in the engineering literature.

In order to better understand the effects of an intervention in a particular school system, systems engineering and education research approaches need to be combined, leveraging system dynamics, agent-based modeling, and social network analysis where appropriate. This view is consistent with Maroulis and Guimerà, who call for the use of complex systems analysis, agent-based modeling, and social network analysis techniques in education policy [26]. These techniques are appropriate when dealing with social systems; people can be modeled as agents or as nodes of a social network connected through links representing their relationships.

2.3 Interventions in the School System

Much research has been conducted on analyzing technology tools and designing interventions for teaching and learning. But, even though a particular intervention was effective in one school setting, it may not necessarily be effectively implemented in another school setting. There has been a shift in the past decade in educational research from designing the interventions to studying the implementation of interventions in the school system. Fixsen [27] presents a detailed synthesis of the implementation research literature marking this shift. A number of studies, falling under the category of *effectiveness research* (ER) within educational research, have been conducted with the aim of identifying interventions that can work in a wide variety of settings [28]. However, treatment effects of interventions vary widely from one school system to another. As mentioned before, there are several factors that can affect the implementation of interventions; some of the factors that have been shown by research in educational interventions to be critical in school-wide change are the following: professional development, leadership, organization and school structure, and resources and support [29-33]. Professional development can be defined as the training given in instructional/pedagogical strategies, content knowledge, and/or new technology tools to the teachers and staff at the school. Leadership involves the role of school administration through various stages of implementation of the intervention. Complementary to leadership is organization and school culture, which might involve teacher autonomy as well as a supportive social network in the school.

Of course, every school and every intervention is different and there can be varying needs based upon the setting in which the particular intervention is being implemented. The above studies try to generalize the factors needed in order to achieve school change. To address this issue, a new research field has emerged in education research called *design-based implementation research* (DBIR) [34]. The aim of DBIR is to simultaneously develop interventions and improve their implementation. The DBIR community calls for research to address the following questions which are not being addressed by ER [35]:

1. How to incorporate considerations of implementations and sustainability in the intervention development?
2. How to change conditions that inhibit implementation of potentially effective interventions?
3. How to promote adaptations of interventions?

The work done in this thesis is at the heart of DBIR and the questions posed above by the DBIR community can be tackled with the help of the ESIM framework. Different interventions have different demands in terms of resources, agents' attribute levels, and relationships. To add to this, even the same types of interventions can lead to different outcomes in different school settings. Hence, there is a need to develop a framework that can analyze *specific* implementations of the interventions given the context of application. To the best of our knowledge, no attempt has been made to quantitatively analyze the implementation of *actual* educational interventions and the change in the school system as a result of implementing the intervention. This thesis addresses this gap in the literature by using industrial engineering, systems engineering, as well as operations research techniques to analyze interventions which actually took place in real schools with real outcomes for comparison and understanding.

2.4 Agent-Based Modeling Frameworks

Recognizing the value of agent-based modeling to model complex systems, researchers have developed various frameworks in their specific fields of application. The benefits of developing a *framework* are: *(i)* It can be applied across different settings; *(ii)* It enables comparison across models; and *(iii)* It facilitates repeatability of results [36]. A framework has been developed for building agent-based social simulation models [37], which is a conceptualization framework based on Ostrom's Institutional Analysis and Development framework [38]. This framework has been developed to help modelers who are not familiar with software development to conceptualize and implement agent-based models. Agent-based modeling frameworks have also been developed for specific fields. Two such widely used frameworks developed for building models for ecological systems are the Overview Design and Development (ODD) protocol and the MR POTATOHEAD framework [39, 40]. ODD is more focused on communication and reimplementation whereas MR POTATOHEAD is developed to enable better comparison between models. Since both of these frameworks have been developed for ecological systems, which usually do not include human agents, these are somewhat restrictive in applications across social science. Another agent-based modeling framework has been developed for civil infrastructure systems which uses hybrid agent-based modeling and system dynamics techniques to analyze infrastructure policies [41]. Each of these frameworks is well suited for its intended application. A comprehensive review of agent-based modeling frameworks across different fields is beyond the scope of this thesis; however, much can be learned about what features should be included in a framework to analyze the education system by studying such frameworks developed in other fields. To the best of our knowledge, there does not exist a modeling framework to analyze interventions in the education system. In the next chapter, the case study used in the development of the modeling framework is presented, along with a model to analyze the implementation of the particular case study.

CHAPTER 3

EWB CASE STUDY FOR FRAMEWORK DEVELOPMENT

In the previous chapter, a literature review was presented which discusses prior research conducted in analyzing interventions in the education system. A large gap was identified in the literature where there is a need to model the implementation of interventions in the school system using quantitative techniques. Industrial engineering, systems engineering, and operations research provide such techniques and methods which can be used in this setting.

In this chapter, we present a case study[2] which was used to begin the development of the ESIM (Education System Intervention Modeling) framework. An actual case study was chosen to develop the framework so that it is grounded in real-world application. In the following sections, first a description of the case study and the reasons for choosing to model it are given. Then the model development process of the case study is presented, along with results and analysis at the end.

3.1 Introduction to EWB

The case study is an *Engineers Without Borders* [42] (EWB) chapter that was implemented in a magnet school setting through a partnership with the Georgia Institute of Technology, Atlanta (Georgia Tech) as part of a National Science Foundation GK-12 grant [43]. This case study was ideal for the development of the framework because it had two different outcomes over two different years, which helped in uncovering insights

[2] This chapter is an extension of our work in [2]

about the success of this intervention. Also, the scale of this intervention was small enough to test the development and application of the ESIM framework. Further details of the EWB case are discussed in *Section 3.2*. Different steps taken during the model development guided the framework design process, and ultimately led to the development of a more-generalized framework which can be applied across many different types of interventions in the school system. The model developed for this particular case study, to analyze the barriers and enablers for the successful implementation of this intervention, is presented next.

3.2 Model to Analyze the EWB Case Study

Developing a model for the EWB case study required collaboration with the lead teacher involved in this intervention, educational researchers including both some who were involved in the intervention and some who were not involved, public policy researchers, and industrial and systems engineering researchers. This inter-disciplinary collaboration was identified as a key feature for developing a model which was both valid and useful in practice. Throughout the model development process presented below, it will be indicated where each discipline played a critical role.

Overview of the intervention: Before going into depth about the model, some critical features of the intervention are discussed which were used in guiding the model development process. This intervention was carried out in a 99% African-American magnet high school. There was a graduate student from the GK-12 grant who worked in the school one day per week along with the lead teacher. A part of this intervention was to start an after-school club where the students would interact with the teacher and the graduate student teaching fellow to learn how to build a solar cooker. Another part of the expectation of the intervention was to send some of the students along with the teacher and the graduate student teaching fellow to Tanzania to install the solar cooker. Two years of this intervention were analyzed. In its first year, the lead teacher, principal, and graduate student teaching fellow were able to amass sufficient financial resources to enable club members and chaperones to travel to Tanzania to implement the solar cooker project. Seven students were sent to Tanzania in the first year along with the lead teacher and the graduate student teaching fellow. While the EWB chapter continued for a second year, the lead teacher and principal of the school had changed. The enthusiasm and support for intervention of the new agents was lower than that of the initial agents. This led to a different outcome, which is analyzed through the model.

<u>Scale</u>: The first step in creating the agent-based model was to determine the scale at which the model would be developed. The scale of analysis for this intervention is of an after-school club. Although this is not a comprehensive school-level intervention with many school teachers and administrators involved, the dynamics of the intervention are still very interesting to model because of interactions taking place between the school agents and community. The social network structure of the agents involved in this intervention is non-hierarchical, with the resource flows/interactions taking place amongst any agents that are linked to each other, i.e., there is no single agent who coordinates the overall system dynamics.

Another dimension of scale which had to be decided was the time horizon and time step size. The time horizon for this intervention is one school year, with the after-school club meeting for the first nine months and the remaining time being for the trip to

Tanzania to implement the solar cooker. Each time step is modeled as having a length of one month. This was decided in consultation with 'educational researchers' to allow appropriate time for an attribute to change. Sensitivity analysis is conducted later by changing the time step.

Agents, attributes and environment: The network of agents for the EWB case study is shown in Figure 1. In this case, the principal, teacher, teaching fellow, and school partner were key agents, as well as the community, which is modeled as one agent rather than considering individual entities. There are two main reasons for modeling the community as a single agent: there is not enough data available to model each community member as an individual agent, and for the purpose of this model an average over the whole community population can be taken to model the behavior of the community as the socioeconomic status and population are reasonable predictors of community behavior in this model. However, within the class of students who participated, we do model individual students as having their own attributes and relationships. In the figure, certain students are shown to have larger or smaller spheres of influence—they may be popular or they may be leaders, which would affect how many other students they influence.

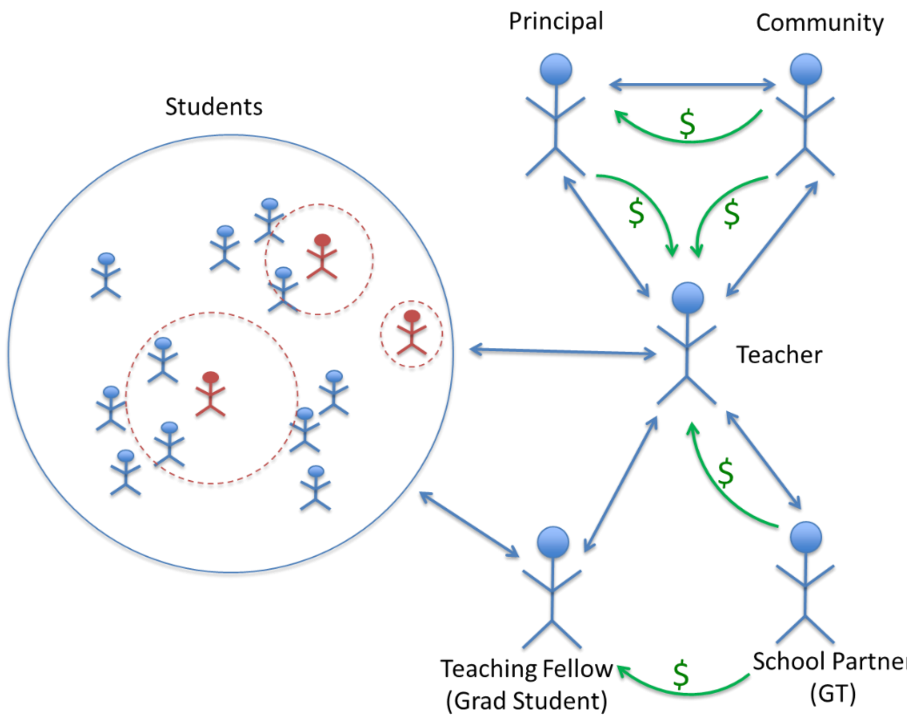


Figure 1: Agent network for EWB case study

It should be noted that this agent network is by no means exhaustive. For example, the federal, state, and local governments that play a major role in the education system could have been included, but these are above the scale of analysis and are treated as constants in the model. This is also reasonable in this case because the government did not directly impact the variables of interest within the short time horizon of this intervention. However, as it will be demonstrated by the varying agent networks in the different case studies modeled in this thesis, different cases require different agent networks.

The next step is to decide which attributes to model for each agent. In the EWB case, the following attributes are modeled for each agent class:

- **Students:** Career aspirations, cognitive ability, self-efficacy, perseverance, socioeconomic status (SES)
- **Teacher:** Support for intervention, cognitive ability, self-efficacy, organizational citizenship
- **Principal:** Support for intervention, self-efficacy
- **Community:** Support for intervention, population, SES
- **Teaching Fellow (Grad Student):** Leadership, cognitive ability, self-efficacy, perseverance
- **School Partner (Georgia Tech):** Support for intervention

Research tools exist for measuring some of these attributes and many others for conducting studies within the school [44]; however, for this case we used expert knowledge from the first year's project lead teacher to estimate some of these attributes on a Likert scale. This might introduce some bias in the data collected because of the involvement of the agent, from whom data is being collected for some attributes, in the intervention itself. That is why sensitivity analysis is important, in order to look at changes in the model outputs as the inputs are varied. Again, the above list of attributes is by no means exhaustive and there is always a tradeoff between model complexity and model accuracy when making such decisions. The list of agents' attributes to model for this case study was decided in consultation with the 'practitioners' and 'educational researchers'.

The environment in this case study includes the relationships between the agents and the exchange of resources. The blue arrows in Figure 1 denote the inter-relationships between the agents, and it is assumed that all of the students are connected via intra-relationships. In this case the inter-relationships are modeled bi-directionally whereas the intra-relationships are modeled symmetrically. This is because the relationships between the students were observed to be mutual whereas the relationships between the other agents depended upon the agent's perspective from which the relationship was being assessed. While unlinked agents may interact with each other, it is assumed that their limited interaction will not impact this case study. The resource flows are represented by the green arrows in Figure 1. In the EWB case study, there is a need for money to send the agents to Tanzania to install a solar cooker. The sources of money in this case are the community and the school partner, and money flows from these agents to the principal and the teacher through the set of links that connect them. The relationships modeled for this case study are those that either facilitated resource flows or those which were directly affected by the intervention.

Agent-based model rules: Each agent that is modeled for the case study has a specific role in the intervention. The teacher and the graduate student teaching fellow interact with the students to teach them how to build a solar cooker, the principal and teacher interact with the community to gather financial support for the intervention, the teacher collaborates with the school partner to raise money for the trip to Tanzania, and so on. Table 2 summarizes the behaviors modeled for the different agent classes in this case study.

Table 2: Agents' behaviors for the EWB case study

Agents	Behaviors
Students	• Learn to build the solar cooker • Develop relationships with other students • Go to Tanzania to install the solar cooker
Teacher	• Teach how to build a solar cooker • Secure funding for the Tanzania trip through the school partner and the community • Go to Tanzania with the students
Teaching fellow	• Teach how to build a solar cooker • Develop collaboration between the teacher and school partner • Go to Tanzania with the students
School partner	• Provide monetary support for the Tanzania trip
Principal	• Allocate school resources to facilitate the EWB club meetings • Secure funding for the Tanzania trip through the community
Community	• Provide monetary support for the Tanzania trip

As a result of the above agent behaviors, there are three major changes happening in the system: the attributes of the agents change, their relationships change, and the resources flow between the agents. Changes in the system state, composed of agents' attributes and relationships, are modeled as a Discrete-Time Markov Chain (DTMC). It is assumed that change can be modeled as taking place in discrete time steps and that changes in the agents' attributes depend only on the current attribute level and current relationships, not on past system states. Markov models have been used before in the modeling of healthcare systems and human interactions [45, 46]. This simplification helps to make the model tractable. To model the DTMC, the time horizon for the intervention is divided into discrete time periods, during which the attributes of the agents and their relationships have some probability of change. There are three possible movements in the states: improving, no change, or worsening. As shown in Equation (1), the state change probability equation is made up of two components: one is internal to the intervention and corresponds to the phenomena captured in the model, that is, an *internal* component; and the other is an *external* component that accounts for parameters external to intervention and outside the scope of the model which may affect the state change.

$$p_{change} = w_{internal} \cdot p_{internal} \cdot f_s + w_{external} \cdot p_{external} \tag{1}$$

where p_{change} is the overall probability of change, $p_{internal}$ captures the 'modeled' aspects of change, and $p_{external}$ captures the aspects that are not modeled but that still can affect the change probability. These probabilities are vector quantities representing the three possible state changes, $[p_{improve}, p_{stay}, p_{worsen}]$, which add up to 1. In addition, there are weights, $w_{internal}$ and $w_{external}$, which quantify the percentage of change probability associated with each of the internal and external probability vectors. These are non-negative and the sum of these weights is always 1. Individual weights can be tuned for each model and their impacts can be investigated using sensitivity analysis, which will be discussed in subsequent sections. Finally, the internal portion of the equation includes a multiplicative factor f_s which captures an 'S' curve pattern in learning. This is used to dampen the probability of change when the attribute level is particularly low or high, in which cases it is harder to change the attribute level. The rules for the three main sub-models are discussed below.

Rules: Change in Attributes

The internal change probability, $p_{internal}$, for a particular agent is a function of three qualities: the current attributes of the agent, the agent's current relationships with other agents, and the attributes of the agents with whom the agent has relationships. Since change is modeled as a DTMC, at each time period, the change in the attribute is a function of the variables in the previous time period. As an illustration, consider the internal change probability for the attribute **'support for intervention'** of the agent **'community'**, shown below:

$$p_{int-comm}(t+1) = w_p \cdot inter_{p-c}(t) \cdot support_p(t) + w_t \cdot inter_{t-c}(t) \cdot support_t(t) \qquad (2)$$

where $p_{int-comm}(t+1)$ is the internal probability of change in the community support for intervention in time period $t+1$ and is dependent only on the system state at time t. As seen in Figure 1, the community has relationship ties with the principal and the teacher. As such, the first term is for the relationship with the principal, where w_p is the weight for this term, $inter_{p-c}(t)$ is the interrelationship between the principal and the community at time t, and $support_p(t)$ is the principal's support for intervention at time

t. Similarly, w_t is the weight for the teacher term, $inter_{t-c}(t)$ is the inter-relationship between the teacher and the community at time t, and $support_t(t)$ is the teacher support for intervention at time t. Again, $w_p + w_t = 1$ and weights may be tuned for the particular case. This equation is representative of attribute change probability equations for different agents in the model.

Rules: Change in Relationships

Two concepts from social network analysis, homophily and structural balance, are used to model the change in relationships of the agents in this case study. The following equation models the internal **intra-relationship** change probability for two **students *i* and *j*** as a function of their structural balance and homophily**:**

$$p_{int-ij}(t+1) = \underbrace{w_{sb} \cdot \operatorname*{avg}_{k \in S \setminus \{i,j\}} \left\{ \operatorname{sign}\left(intra_{ik}(t) \cdot intra_{kj}(t)\right) \cdot \min\{|intra_{ik}(t)|, |intra_{kj}(t)|\} \right\}}_{\text{structural balance}} + \underbrace{w_h \cdot h_{ij}(t)}_{\text{homophily}} \tag{3}$$

where $p_{int-ij}(t+1))$ is the internal change probability in the relationship between student i and j in time period $t+1$, and S is the set of all agents belonging to the student population. To model structural balance, the weight for this term, w_{sb}, is multiplied by the average relationship strength for all the students in S with whom *i* and *j* interact. The sign function captures the structural balance tenets described previously, while the minimum function depicts that the magnitude of the effect is constrained by the *weaker* of the two relationships that students i and j have with another student, k. Therefore,

$intra_{ik}(t)$ is the intra-relationship between students i and k and $intra_{kj}(t)$ is the intra-relationship between students k and j in time period t. For the homophily component, w_h is the weight and h_{ij} is the measured level of homophily between students i and j. Again, $w_{sb} + w_h = 1$ and these weights may be adjusted appropriately. The change in the relationships between the other agents is modeled similarly; however, relationships between different agent classes are bi-directional, meaning that the relationships do not have to be mutual.

__Rules: Resource Flows__

In the EWB case, the resource is money to send a group to Tanzania, and this demand drives the school partner and the community to provide funding. Each agent has a budget which cannot be exceeded. For this model, it is assumed that the school partner had a specific budget from a grant and that the community budget depends on its tax base (as an indicator of the socio-economic status of the community) and the number of people in the community. However, the full budget amount is not always supplied; money flows are dependent on the strength of the relationship between the agents through which the money must move. Equation (4) represents the flow of money at time $t + 1$ from the school partner to the teacher:

$$m_{sp\to t}(t+1) = inter_{sp,t}(t) \cdot support_{sp}(t) \cdot b_{sp} \tag{4}$$

where $m_{sp\to t}(t+1)$ is flow of money from the school partner to the teacher in time period $t + 1$, $inter_{sp,t}(t)$ is the interrelationship between the school partner and the teacher in time period t, $support_{sp}(t)$ is the school partner's support for intervention at time t, and b_{sp} is the available budget for the school partner. Other money flows are modeled similarly. This completes the set of rules which govern the state changes in the agent-based model.

<u>Conceptual model validation</u>: The conceptual model for the EWB case is complete now. At this stage the conceptual model built can be verified using subject matter experts. There were a total of 8 subject matter experts who were used to validate the conceptual model. These were ‘practitioners’, ‘educational researchers’, and ‘public policy researchers’. This resulted in refining/modifying the agents’ attributes and change equations. The conceptual model presented above is the result of a number of iterations of change with the guidance of the subject matter experts.

<u>Criteria for intervention success</u>: For the purpose of analyzing this school intervention, the intervention success criteria were composed of attributes and resources that would enable the occurrence of the trip to Tanzania to implement the solar cooker. This translates into the following criteria: there should be a minimum of four students that are sent to Tanzania so that the solar cooker can be installed within the limited time the students are there; these students should have positive relationships amongst themselves, the teacher, and graduate student teaching fellow who will be sent along with them; the test scores of these students should be above a certain threshold; and finally there should be enough funding to send the students, teacher and teaching fellow for this trip, which translates to ~$20,000. In this case study, since the trip to Tanzania happened in one year and did not take place the following year, the above success criteria were used to help in understanding this difference in the outcomes.

Data collection: To simulate the case study, the initial state of the system must be defined. For the EWB case, data were collected regarding the initial and end states of the system with the help of a teacher involved in the EWB reform—this is the same teacher agent shown in Figure 1. Even though some data were available for the students involved in the intervention, for the most part teacher knowledge was assumed to be sufficient. Because this was the first test case for both developing and applying the framework, relying on expert knowledge was considered to be a viable option. In the SLIDER case study presented in Chapter 5, more-robust data collection instruments are used.

Computer simulation model: The agent-based simulation model was built using the object-oriented programming language *C#* in Microsoft Visual Studio 2010. The system configuration was Windows 2007, 64 bits, 8GB Ram, and 1.73 GHz Processor. An object-oriented programming language was chosen because it allowed the agent classes to be modeled as different classes in the program and agents as objects in the program. It also provided the flexibility to implement the various agent-based model rules developed for this intervention.

3.3 Results and Analyses

Simulation results: Since this is a stochastic model, first the internal variance of the model is estimated, and then based upon that estimate, sufficiently large number of runs are made to analyze the outcomes for the two years in order to ensure that the internal variance is low. Table 3 shows the variance in the various output variables of interest as well as the running times for the number of runs. The support for the intervention attribute of the community is quantified using a scale from -2 to 2 and money is measured in dollars.

Table 3: Mean and standard deviation of model outputs

No. of samples	No. of iterations	Community Support		Money through Principal ($)		Money through Teacher ($)		Running time (mins)
		Mean	Std_dev	Mean	Std_dev	Mean	Std_dev	
100	100	1.83	0.02	26395	95.84	17696	48.16	10
1000	100	1.83	0.02	26401	87.86	17699	44.10	88
100	1000	1.83	0.01	26401	27.21	17699	13.72	90

In the above table, the number of iterations represents the number of runs made, over which an average of the results is taken. This is done so that the internal stochasticity in the model results can be reduced to treat the results as almost deterministic. For example if the number of iterations is 100 then these 100 stochastic iterations constitute a single almost-deterministic iteration which is the average of those 100 iterations. Next, to estimate the stochasticity remaining in these almost deterministic model runs, a sample is taken whose size represents the number of runs of such almost deterministic model iterations. As the number of iterations increases, the standard error of the estimates decreases. The number of samples represents the number of runs made to calculate the standard deviation in the model when an average is taken over the number of iterations. Table 3 shows that the variance in the model, when an average of the results over 1000 model iterations is taken, is three orders of magnitude lower compared to the mean. Since this was a sufficiently low variance, for the purpose of analysis whenever an input variable or parameter is changed, 1000 iterations of the model will be run and the average over those will be used to estimate the change in the output. Now the results of year 1 and year 2 can be discussed and compared.

Year 1 Results: The final state at the end of the simulation for the first year met all the criteria for the trip to Tanzania to be feasible. Specifically, the community support for intervention reached was 1.8 on a scale of (-2, 2), the total money generated to meet the demand was about $44,100, and the relationship criteria were satisfied. Since more money was raised than the minimum requirement of $20,000, more than 4 students were sent to Tanzania. Hence, the end state reached by the model was consistent with reality, where 7 students, the teacher, and teaching fellow successfully installed a solar cooker in Tanzania.

Year 2 Results: In the second year of the intervention the social network changed drastically. The new principal and lead teacher at the school did not have strong community ties. In this case, the final simulation state did not meet all the criteria for the trip to be feasible. Even though there were qualified students with positive relationships, the community support for intervention was not very high, and so the estimated total money generated was only about $10,400, which is significantly less than the required $20,000. This end state is also consistent with reality in year 2 when the trip did not take place due to the shortage of funds.

Sensitivity analysis: Sensitivity analysis was conducted with respect to the weights used in the change equations, parameters of the S-curve, and the granularity of time periods. Table 4 shows the effects of change in the principal and teacher weights from Equation (2) on the community support for intervention and total money generated. A few extreme values are shown on either end of the spectrum, and the weights are incremented in step sizes of 2% in the range of interest. These weights represent the relative influence of the teacher and principal on the community's attributes; the total weight must add up to 100%, and for example, a weight of 0% for the teacher implies that the community is only affected by the principal. The range of interest was determined using expert judgment from the teacher involved in the intervention. The support for intervention attribute is quantified using a scale from -2 to 2. The negative scores on this scale represent negative traits of the attribute, with zero being neutral and positive numbers indicating positive traits. The actual numbers on this scale are unimportant—it is the relative change in the attribute levels across time periods that defines the success or failure of the intervention. The principal and teacher initial support for intervention are taken as 1 and 2, respectively. This is different from the starting conditions in year 1 when both the teacher and principal support for intervention were 2. This change is made to better conduct the sensitivity analysis—the starting conditions in year 1 would have masked the effect of change in these weights on the final state. The community support for intervention and the total money shown is the level reached by the end of the time horizon for this case study and is based on the average of 1000 simulation runs. In each simulation run, the agent-based model is simulated over the duration of the intervention. At every time step, there is a probability of change in the state of any attribute, given by the probability vector [$p_{improve}, p_{stay}, p_{worsen}$]. To decide whether or not the attribute state changes, a uniformly distributed random variable (U) is generated between [0,1] and its value is compared to the change probability vector as follows: (a) If, $U < p_{improve}$, then the attribute level increases by the step size; (b) Else if, $U \geq p_{improve}$ & $U < (p_{improve} + p_{stay})$, then the attribute level stays the same; (c) Else, the attribute level decreases by the step size. A new value of U is generated at each time step, for each attribute. Since, it is a probabilistic model, where the same set of inputs can lead to different outputs, an average of 1000 simulation runs is taken to analyze the results.

Table 4: Sensitivity analysis with respect to weights

No. of iterations	**Principal Support t_0**	**Teacher Support t_0**	**w principal**	**w teacher**	**Community Support t_n**	**Total Money t_n**
1000	1	2	0%	100%	1.83	44100
1000	1	2	30%	70%	1.59	36406
1000	1	2	50%	50%	1.41	30315
1000	**1**	**2**	**60%**	**40%**	**1.32**	**27373**
1000	**1**	**2**	**62%**	**38%**	**1.30**	**26652**
1000	**1**	**2**	**64%**	**36%**	**1.29**	**26251**
1000	**1**	**2**	**66%**	**34%**	**1.27**	**25667**
1000	**1**	**2**	**68%**	**32%**	**1.24**	**24834**
1000	**1**	**2**	**70%**	**30%**	**1.24**	**24453**
1000	1	2	80%	20%	1.14	21627
1000	1	2	90%	10%	1.05	19003
1000	1	2	100%	0%	0.93	16184

There are two key observations that can be made from this analysis. One is that the community support for intervention and the total money generated are indeed responsive to the change in weights; specifically, the final states of these attributes decrease as the weight for the principal increases because the principal support for intervention is less than the teacher support for intervention. Another observation is that for the boldfaced portion of the table, where the weights are assumed to lie for this case study, the weights do not have a significant impact on the outcomes. This is a good sign, as we are interested in determining whether or not the final state is likely to fall under an umbrella of acceptable states without knowing this weight precisely. In this case, all the final monies exceed the required $20,000. In addition to this example, sensitivity analysis was also carried out for the other weights, parameters of the S-curve, and the number of time periods. Tables 5 and 6 below show the effect of change in the number of time periods and the S-curve parameter on the model results, respectively. The model is again sensitive to these parameters, but not to an extent that the final state would become unacceptable due to slight changes in these parameters. This is useful in establishing the reliability of the model results. As before, the results presented are for an average of 1000 simulation runs.

Table 5: Sensitivity analysis with respect to the number of time periods

Number of time periods	**Step size**	**Community Support**	**Money through Principal**	**Money through Teacher**
17	0.12	1.8	26609.2	17804.7
10	0.20	1.8	26527.4	17762.9
9	0.22	1.8	26421.7	17708.9
8	0.25	1.8	26398.6	17698.7
7	0.29	1.8	26383.3	17698.0
3	0.67	1.8	25751.8	17363.8

In Table 5, step size is the magnitude by which the state of a variable changes and is calculated using the formula given in the model design phase of the framework. As the number of time periods in which the time horizon is divided decreases, the length of a time step increases and so does the step size. The effect of using a different number of time periods does not change the model results much because the assumption is that an attribute can go from zero to its maximum or minimum level within the time horizon of this intervention. However, for other case studies for which this assumption does not hold, using a different number of time periods might have a larger impact on the output. The standard error in the above results is two orders of magnitude lower than the mean, based upon 1000 simulation runs.

Table 6: Sensitivity analysis with respect to S-curve parameter

S-curve parameter	**Student 1 Test Score_t_0**	**Student 1 Test Score_t_n**
0.00	-1.00	-0.84
0.05	-1.00	-0.85
0.10	-1.00	-0.87
0.15	-1.00	-0.90
0.20	-1.00	-0.88
0.25	-1.00	-0.90
0.30	-1.00	-0.89
0.35	-1.00	-0.91
0.40	-1.00	-0.94

In Table 6, the S-curve parameter represents the magnitude of dampening due to the S-curve factor f_S in Equation (1). As the S-curve parameter increases, the magnitude of dampening increases. That is why, when the S-curve parameter is at its maximum, the change in the student test scores is at its minimum. When this parameter is zero, there is no role of f_S in the change equation. As the value of this parameter is increased, the impact of f_S increases. In Table 6, it can be seen that for student 1 the initial test score level is -1 on a scale of -2 to 2, which is a low magnitude state. Hence, when f_S does not play any role the improvement in the test score is higher and as f_S starts to impact change more, this improvement in the test score decreases. This happens because the initial state is low and the S-curve behavior dampens the upward movement when the state of the attribute is very low. Hence, we see from the sensitivity analysis conducted for various parameters that the model's outputs are responsive to the change in these parameters; but not to an extent that small variations in these parameters lead to drastically different results. This completes the sensitivity analysis for this case study.

Method of Morris: After sensitivity analysis, we used the Method of Morris (MoM) [47] to determine a subset of input variables, from amongst the larger set, which have a significant impact on the outcome of interest (response variable). This is explained in Chapter 4 in more detail. The MoM experiment is conducted with 24 different input variables, including attributes of the teacher, principal, and students, and the relationships between them at the start of the intervention. The response variable used to analyze the effect of the inputs is the total money generated by the end of the intervention. Figure 2 depicts a graph of the mean and standard deviation of the money generated due to each input variable. The mean of the effects is the x-axis and the standard deviation of the effects is the y-axis. The dotted lines correspond to the equation: Mean = $\pm 2 \cdot$SEM, where SEM is the standard error of the mean and is equal to the standard deviation divided by the square root of number of random orientations for each input, which for this analysis is taken as 10. The circles represent the different inputs. Inputs lying outside of the 'v' or far from zero have an impact on the response variable that is statistically different than zero.

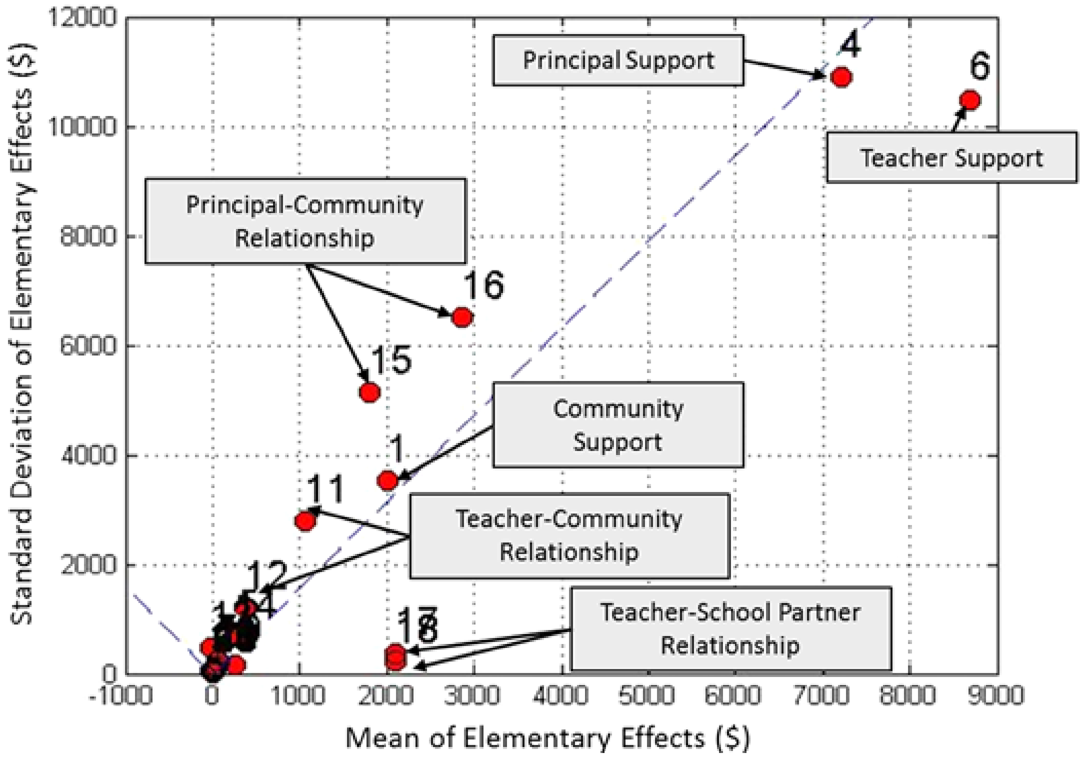


Figure 2: Method of Morris, total money generated

From the figure we see that, the teacher, the principal and the community support for intervention, the bi-directional relationships between the teacher and the community, the bi-directional relationships between the principal and the community, and the bi-directional relationships between the school partner and the teacher has all have a significant impact on the money generated for the Tanzania trip. Clearly, if the success criteria were different, different attributes may dominate the response. MoM is versatile enough to accommodate these changes and isolate the main effects in a large collection of variables, though it requires the user to define the success criteria correctly.

3.4 Validation

The following steps were completed to validate the model results for the EWB case. For larger case studies, like the one presented in Chapter 5, a more-comprehensive validation process is used. Note also that the validation step was not just carried out at the end of the modeling process, but throughout the model development process.

<u>Conceptual model and face validation</u>: In order to complete this step, 8 subject-matter experts were used to test whether the model and its behavior were conceptually logical and whether the model's input-output relationships were reasonable. These subject-matter experts were school teachers, educational researchers, and public policy researchers. They tested the model for completeness, consistency, coherence, and correctness as described in the *4C's* framework during the model validation phase of the framework in Chapter 4. This step was also implemented immediately after the conceptual model was developed, before developing the computer simulation model.

<u>Degenerate and extreme condition tests</u>: While analyzing the model results via sensitivity analysis and the Method of Morris, we analyzed the effect of different combinations of values of input variables and parameters on the model output. This helps in testing the behavior of the model under extreme conditions and different ranges of inputs. The behavior of model as seen from the model analysis section is consistent, plausible with the change in inputs as well as under extreme and unlikely combinations of inputs and parameters of the system.

Event validation: The occurrence of events, which in this case was only the 'trip to Tanzania', was compared to reality. For both years the occurrence or non-occurrence of this event was consistent with the reality.

Internal validation: Several evaluations of the model were used to determine the internal stochastic variability in the model. The standard deviation in the results for 1000 model iterations was three orders of magnitudes less than the mean, which implies model consistency.

Data validation: For this case, the data validation required comparing predicted money generated using the model to actual money generated in reality. In addition, attribute levels were verified for accuracy using a subject-matter expert. For larger and more-comprehensive cases, more-comprehensive data validation would need to be performed.

Parameter variability - Sensitivity analysis: The model was run under different sets of parameter and input conditions and the output was analyzed. This was discussed before in the *Results and Analyses* section.

This completes the model development and analysis of the EWB case. The process involved building an agent-based simulation model, analyzing the results, as well as validating the results. With the help of this process, we created a generalized framework, applicable to a larger set of interventions in the education system. This framework is presented in the next chapter.

CHAPTER 4

MODELING FRAMEWORK: ESIM

In the previous chapter, the model developed for analyzing the *Engineers Without Borders* intervention was presented. The modeling process undertaken during this case study helped in the creation of the modeling framework presented in this chapter. The framework contains the list of steps to implement and the methods to use while building a model for a particular school intervention.

4.1 Introduction to ESIM

The ESIM (Education System Intervention Modeling) framework was developed to analyze a broad range of interventions in the K-12 school system. Collaboration across different disciplines such as public policy, education research, industrial and systems engineering, and practitioners is critical for developing a comprehensive and validated framework that can be applied in practice. Figure 3 gives an overview of the roles that each discipline played in both the development and application of the framework.

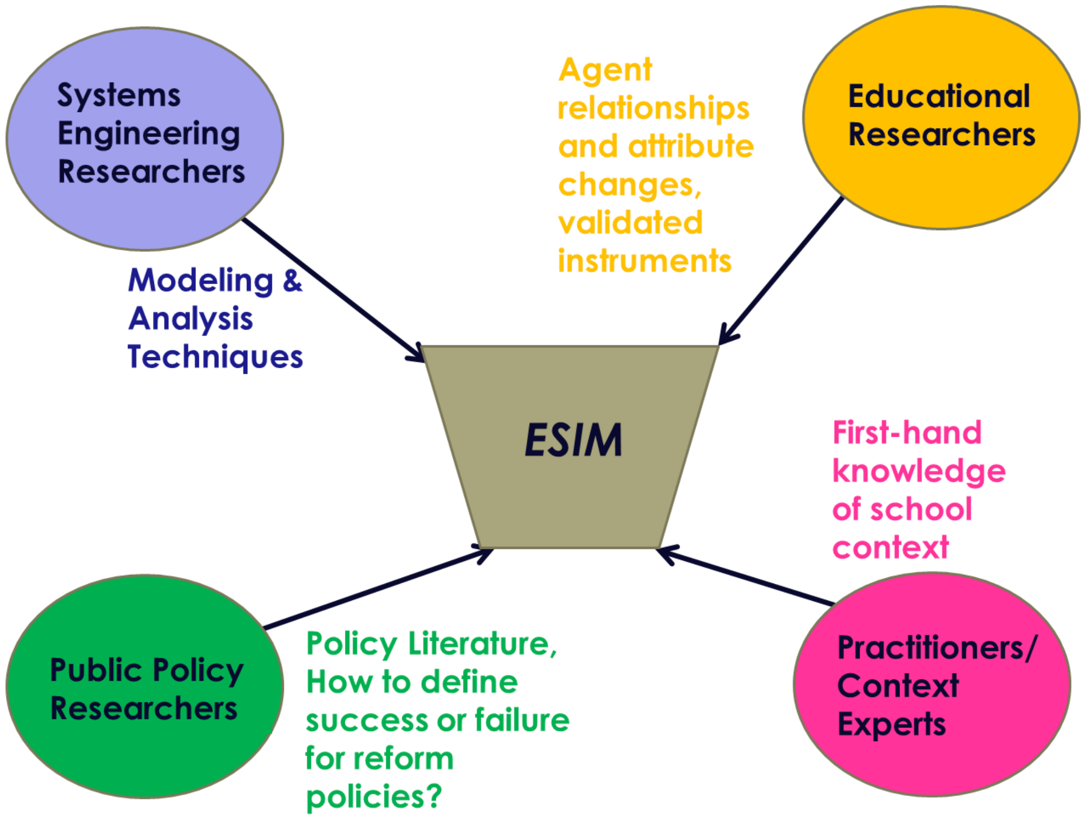


Figure 3: Inter-disciplinary collaboration for framework development

Systems engineering and operations research is at the heart of the framework that brings all of the other disciplines together. Techniques from systems engineering include agent-based modeling and social network analysis, as discussed in previous chapters. But in order to develop such models, knowledge about the agents, attributes, and relationships to be modeled is required and this is provided by the 'educational researchers'. The 'public policy researchers', on the other hand, provide the knowledge about the reform policies and what defines a successful or unsuccessful school intervention. Also, since the framework is being applied to build models for school interventions, it is equally important to take inputs from the 'practitioners', who are school teachers and reformers, to get first-hand knowledge of the context in which the intervention is taking place. Once the framework is developed, the following policy cycle takes place, as described in Figure 4:

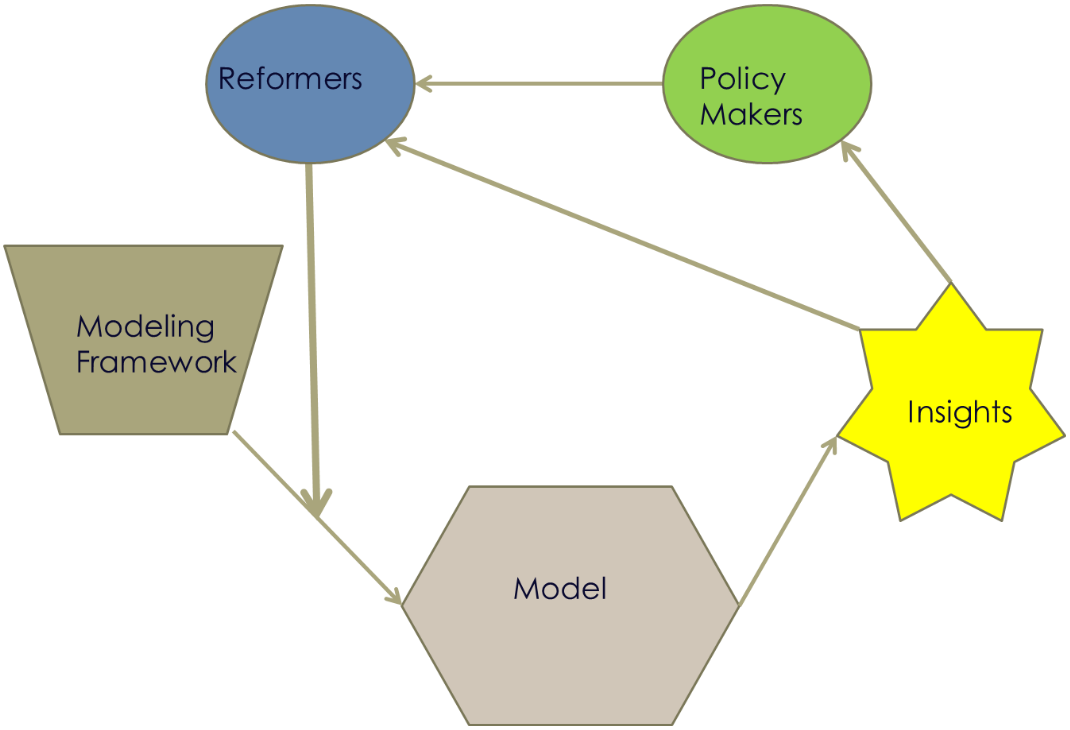


Figure 4: Policy cycle as a consequence of applying the framework

As shown in the figure, when the framework is applied to develop a model for a particular school intervention, insights will be gained about the barriers and enablers of a successful implementation of the intervention. Such insights provide feedback to the policy makers and reformers who could then adapt the intervention accordingly and update the model.

In the remainder of this chapter, the modeling framework is described in detail. The use of the framework to build a model in a particular case is distributed across four phases: model definition, model design, model analysis, and model validation. The framework is applied starting with the model definition; however, the framework application is not linear, and many phases are revisited throughout the process. Figure 5 shows the modeling cycle as the framework is applied.

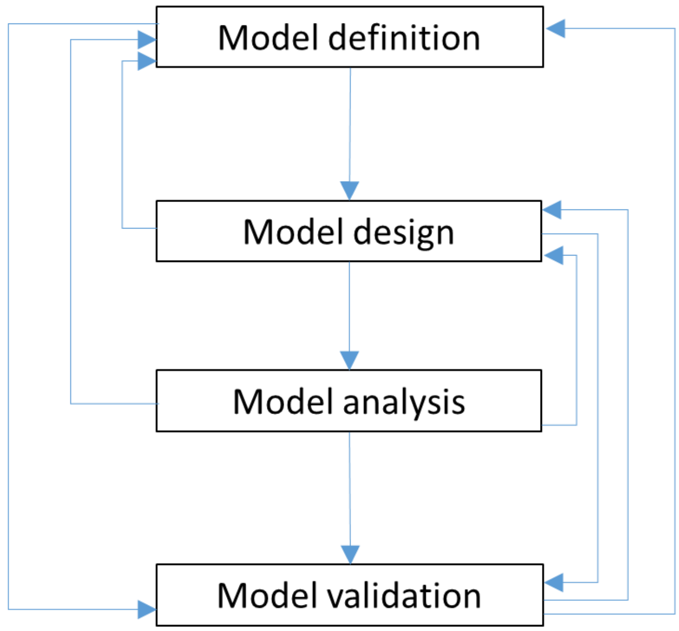


Figure 5: Modeling cycle

The four phases shown above interact with each other throughout the modeling process. The model validation phase is applied after the model analysis phase as well as after the model definition and design phase. The model analysis phase also feeds into the model definition and design phases. Most of the frameworks discussed in Chapter 2 just have a step-wise implementation of different phases of the modeling process, but it is critical to have these phases intertwined in order to build a comprehensive, valid, and useful model. For example, model validation is not something which should be undertaken only at the end of a project. When the conceptual model is built, it should be validated and verified with the help of subject-matter experts. Similarly, when the computer simulation model is built, it should be verified before using its results. Each of the four phases is described in detail in the next section.

4.2 Model Definition

In the model definition phase, a description of the intervention to be modeled is first documented. Then, detailed descriptions about the agents, attributes, and the environment being modeled are provided. Other modeling decisions such as scale and time horizons, are also made in this phase. The following steps address these issues in detail.

Problem definition: The framework developed is used to build models to analyze interventions in a particular school setting. In this task, an overview of the intervention should be provided which includes the description and the objective(s) of the intervention. A clear problem statement helps in understanding the scale at which the intervention is being applied and the variables involved in modeling the particular intervention [48].

Scale of analysis: Determining the scale at which to model is very important when modeling complex systems. The education system from a top down view has a hierarchical structure as shown in Figure 6.

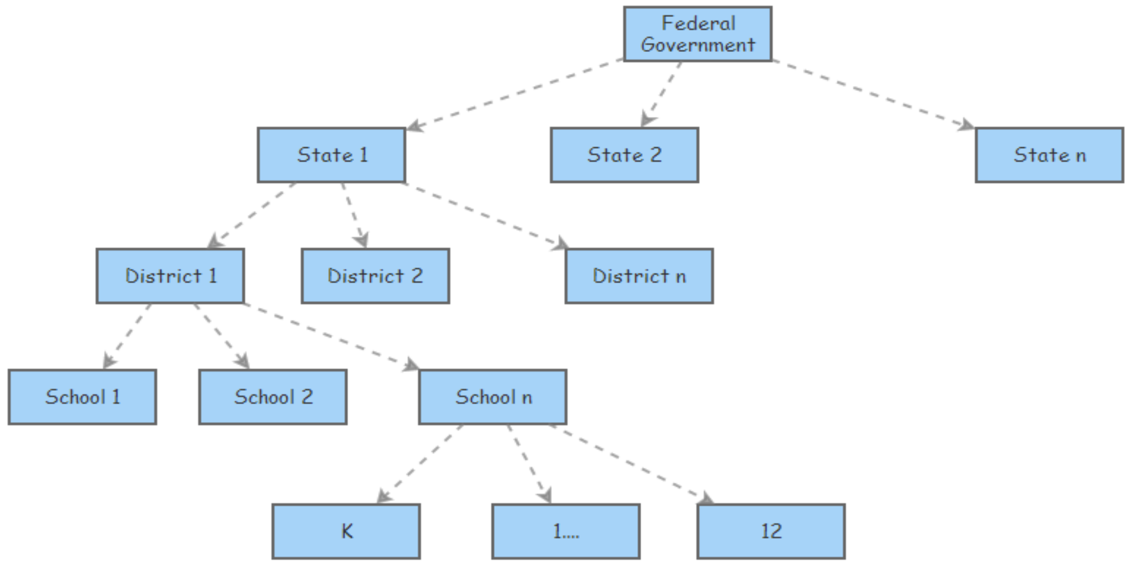


Figure 6: Hierarchical structure of the education system

In the above figure, the lower-most level represents grades K-12. However, a bottom-up view of the education system can look completely different with the hierarchical structure breaking down and getting converted into a network model where there is a great deal more inter-dependence among the parts of the system. Depending upon the scale of the intervention being modeled, different structures arise. When modeling at the scale of the school, the attributes of the agents *above* this scale would be held constant and the attributes of the agents below this scale would be taken as an average [49]. The same principle applies while modeling at any other scale. An example of this is shown in Figure 7, through a simple model at the school level.

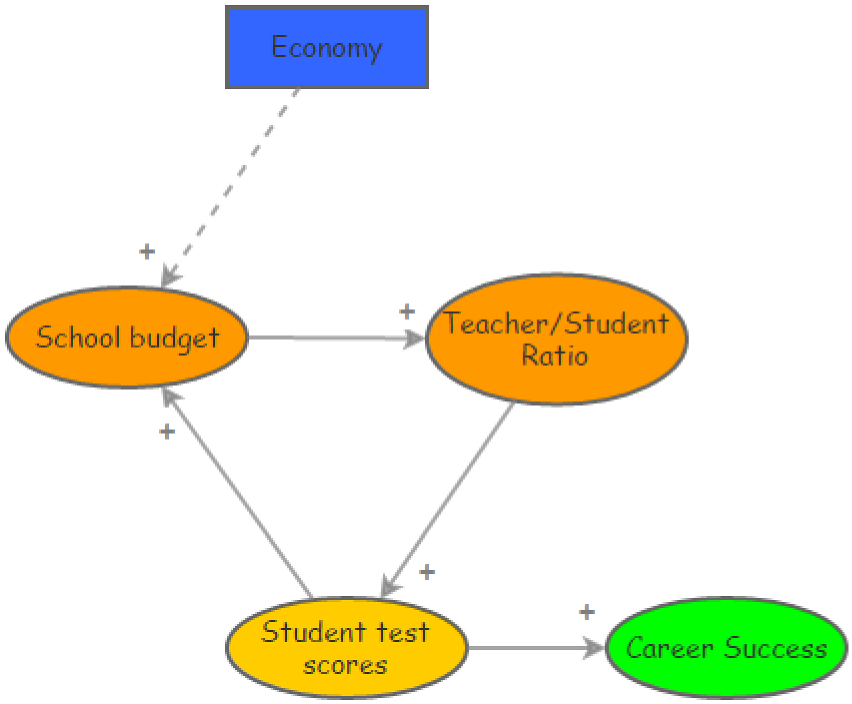


Figure 7: Modeling at the scale of a school

In this model at the school level, the economy, which is at a higher scale compared to the other variables, is an input parameter that is not affected by the dynamics of the model. The variables at lower scales, such as student test scores, are taken as averages for the entire school. In this example, career success of the students is the output of interest being affected by the dynamics of the model. The above model is just for illustrating the role of scale in modeling the agents and their attributes in the intervention.

The scale of the model also helps in deciding which agents to aggregate and model separately. As an example, for a district-wide intervention, the agents in a particular school could be aggregated together as a single agent in the form of the school, whereas for a classroom-level intervention, the teachers and students might be modeled as individual agents. More discussion about the aggregation of agents is provided in the 'agents, attributes, and environment' step of model definition.

Another feature in the framework is to model the same intervention using different scales. For example, if an intervention is being implemented across multiple schools, then implementation at each school could be modeled separately; but the schools can also be modeled together to understand the resource allocations being made across the different schools. The scale at which the intervention is being implemented should be clearly identified to set the right boundaries for the model.

Temporal scale: The next step is to define the temporal scales of the model, which are the time horizon and time step size. The time horizon of the model should match the duration of the intervention being modeled. The time horizon of the intervention plays a critical role in determining which agents and attributes should be held constant, and which should be allowed to vary. Similar to the scale of analysis discussed previously, changes happening at a scale longer than the temporal scale of the model would be held constant, and changes happening at a finer scale than the temporal scale of the model would be taken as an average. Table 7 summarizes the effects of the scale of analysis and temporal scale on modeling change in the system.

Table 7: Effects of scale on modeling change in the system

	Scale of analysis	**Temporal scale**
Larger than scale changes	Constant	Constant
Finer than scale changes	Average	Average
At scale changes	Dynamic/varying	Dynamic/varying

Apart from the time horizon, another important decision is that of determining the time step size of the model. Since the agent-based simulation model being built has discrete time steps, it is important to consider the length of each time step. In order to make this decision, first the scale of analysis has to be identified along with the time horizon. Then, depending upon the agents, attributes, and environment being modeled, a reasonable time step should be chosen for the model that is consistent with the change in the system being modeled. The decision about the time step should be made after discussion with the 'educational researchers'. Again, it should be realized that no single time step may be ideal and sensitivity analysis can be done on the use of different time steps, as discussed in the model analysis section of the framework. This is one of the many instances where a link gets created from model analysis to model definition, as shown in Figure 5. Another feature in the framework is that the time steps can vary within the time horizon. For example, a year can be divided into monthly time steps but certain months where the frequency of agent interactions is high can also be further divided into weekly time steps; see Chapter 5 for an example of where this is exploited.

<u>Agents, attributes, and environment</u>: Once the scale of analysis and temporal scales of the model are chosen, the next step is to examine the network of people who play a role in the success of the intervention. In agent-based modeling, agents simply represent people or classes of people with different attributes. Attributes are characteristics of the agents. To identify the agents involved in the intervention, it is important to work with the 'practitioners' of the intervention at the school level. Some of the agents that are typically modeled for a school intervention are the students, teachers, principal, and other administrators, school partners, local community, and local governmental agencies.

Agents can be modeled as separate people or as a group of people depending upon the purpose of the model and the granularity in the available data. The scale of the model can also help in deciding whether to model different types of individuals as separate agents or group them in a category to model them as a single agent. It should be kept in mind that in modeling any system, there is a trade-off between model complexity and model accuracy; hence, it might not always be possible or efficient to model each individual as a separate agent. To identify which attributes of the agents to model, it is critical to work with the 'educational researchers'. Some of the attributes that are typically modeled are the socio-economic status of the community, test scores of students, content knowledge of the teachers, and leadership quality of the principal. Llewellyn et al. [50] provide a detailed description about the types of agents and their attributes that can be modeled for the K-12 education system. The following tables from that paper give an overview of that description.

Table 8: Agents and their attributes

Entities (Systems of Agents)	Attributes				
	Affective (emotions)	**Cognitive (intelligence)**	**Conative (impulse, volition)**	**Intra-group relationships**	**Inter-group relationships**
Student Population	Morale, motivation, self-expectations	Content knowledge, range of skills, language barriers, students with disabilities	Willingness to work, willingness to take initiative, perseverance, grit	Multiple populations, student culture (cohesiveness vs. divisiveness)	Home life, work, mobility
Teacher Population	Morale, approach to teaching, willingness to learn new ideas	Content knowledge, ability to learn	Willingness to take action, grit	Teamwork, collaboration, planning, trust, communication	Interactions with students, parents, school & system administrators, community
School Leadership	Leadership ability, presence, ethics	Project management & implementation ability, evaluation approaches, budgeting, planning	Problem-solving initiative & ability, confrontation approaches, willingness to take action	Collaboration within the leadership team	Managing up & down--to school system administrators, to teachers, to students, and to community

Table 8 continued

School System Administration	Perspectives on education, political & philosophical leanings, ethics	Management, budgeting, planning, evaluation approaches, testing schedule and philosophy	Micro- vs. macro management, willingness to confront and take action	Cohesiveness, teamwork, alignment	Managing in & out--to schools, to community, to government
Community	Expectations of academic achievement, political & philosophical leanings	Tax base, Socio-economic status	Activism regarding education, impact on school (local media, parent groups, etc.)	Cohesiveness of community, support for school, teachers known by community members	Access to resources (businesses, colleges, etc.)
Government (State DoE, Federal, County)	Political & philosophical leanings, ethics, expectations, sense of urgency	Standards, testing requirements and consequences	Imposition of rules, carrots, sticks, and pressure	Alignment of policies, rules, laws, philosophies	Input from other actors

Table 9: Teacher attribute assessment rubric

Domain	Destructive	Absent	Situational	Constructive
Affective	• Distrust of new ideas and approaches • Caustic mindset • Resistant	• Skeptical of new ideas and approaches • Fixed mindset • Apathetic	• Willingness to learn new ideas and approaches • Variable mindset • Compliant	• Enthusiastic to learn new ideas and approaches • Growth mindset • Committed
Cognitive	• Lack of necessary content knowledge • Extreme difficulty learning new content • Creates learning misconceptions	• Content knowledge at lowest levels of Bloom's taxonomy • Can learn new content given enough time • Perpetuates learning misconceptions	• Content knowledge at middle levels of Bloom's taxonomy • Can learn new content readily • Can identify learning misconceptions	• Content knowledge at top levels of Bloom's taxonomy • Researches and learns new content readily • Transforms learning misconceptions to appropriate learning formations
Conative	• Refusal to take action • Avoids challenges • Undermines action plans & implementation	• Must be persuaded to take action • Immobilized by challenges • Minimal or no	• Willingness to take action • Examines challenges • Involved in action planning	• Enthusiastic to take action • Inspired by challenges • Immersed in action planning

Table 9 continued

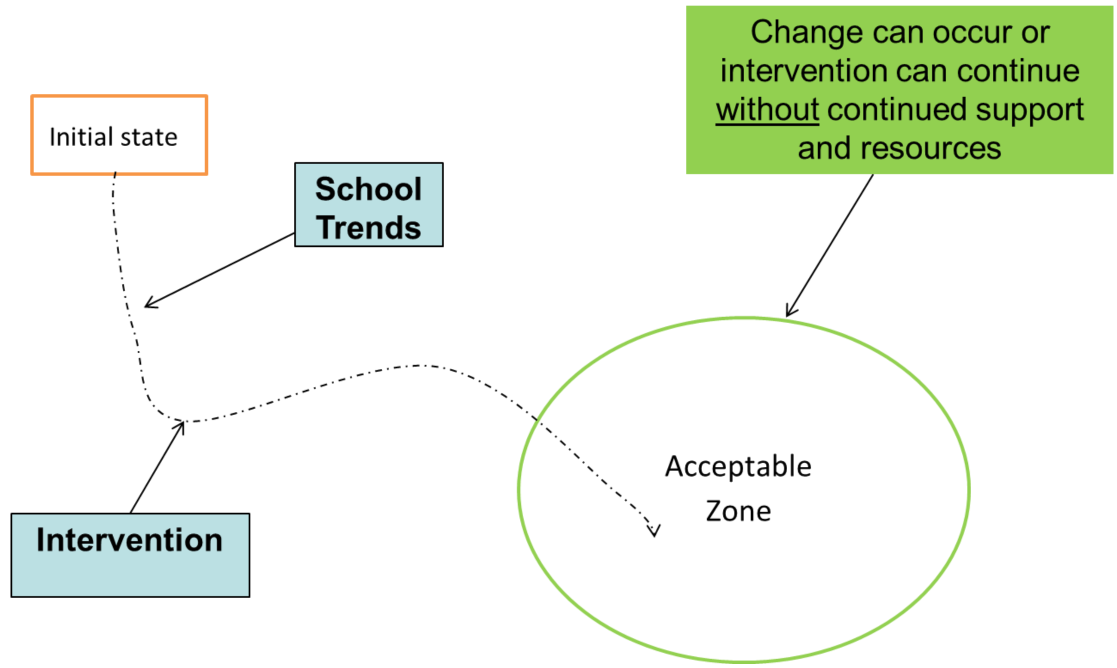
		involvement in action plans & implementation	& implementation	& implementation
Intra-group relationships	• Culture of cynicism • Negative communication • Self-segregated from the group	• Culture of suspicion • Formal or no communication • Isolated from group	• Culture of cliques • Routine communication • Share ideas, resources, and decisions within the clique	• Culture of trust • Frequent prioritized communication • Share ideas, resources, and decisions with all of the group
Intergroup relationships	• Culture of cynicis • Negative communication • Distanced	• Culture of suspicion • Formal or no communication • Detached	• Culture of cliques • Routine communication • Associated	• Culture of trust • Frequent prioritized communication • Aligned

Table 8 provides a detailed list of the agents and their attributes, and Table 9 provides a potential rubric for assessing the teachers' attributes. In Table 8, different agents, such as students, teachers, administration, and government, have attributes that are divided into five categories represented by: affective, cognitive, conative, intra-group relationships, and inter-group relationships. The table provides a description of the type of attributes that would fall under each category for each agent. The term intra-group relationship is used to represent the relationships existing amongst the same agent type, and inter-group relationship is used to represent the relationships existing between different agent types. Table 9 further looks at the five attribute categories for the teacher and presents a possible rubric to assess the various attributes falling under each category. This list is by no means exhaustive, and the agents and attributes to be modeled depend upon the intervention being studied. Research tools exist for measuring some of the above attributes and many others that are germane to school studies [44]. Extensive research has also been conducted to study and measure teachers' inquiry-related attributes [51], which can be utilized while modeling inquiry-based interventions. However, reliable measures may not exist for every attribute. In such cases, there will exist the need for educational researchers to develop measurement tools for those attributes which are found to be important for the success of the particular intervention.

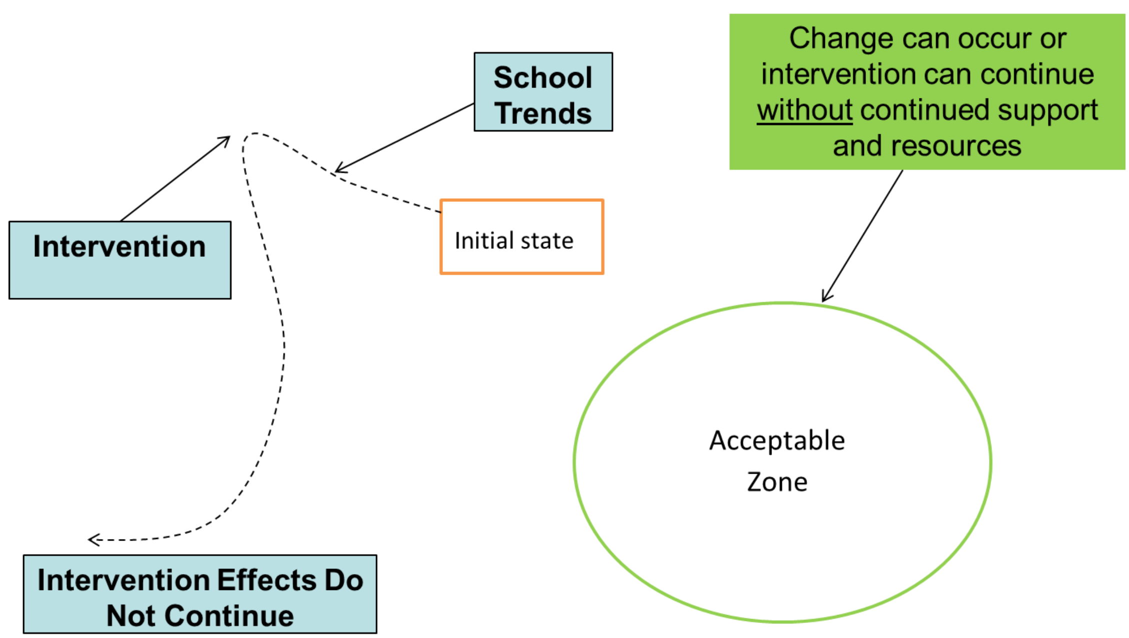
In the agent-based model of a school intervention, the environment includes the relationships between agents and the flow of resources. Flow of resources represents the transfer of money, information, or any other precious entity depending upon the case being modeled. Relationships enable agents to impact other agents' attributes, and each relationship is characterized by an attribute representing the relationship strength. In general, we refer to relationships between different agent classes as inter-relationships, while we refer to the relationships among the same agent class, for example, the student population, as intra-relationships. The relationships between agents can either be modeled as symmetric or bi-directional depending upon the context. Modeling the relationship as symmetric implies that the relationship between those agents is completely mutual. This might not always be the case in the education system, where the relationships can be bi-directional depending upon the agents being modeled. An example of such a relationship could be the relationship between the principal and the teachers in the school. However, when it is not critical to model the relationships as bi-directional, the simpler route of modeling it as a symmetric relationship should be taken. The case study presented in Chapter 3 illustrated how some relationships are modeled as bi-directional and others as symmetric. Finally, the agents, attributes, and environment being modeled are dependent upon both the scale of analysis and the temporal scale at which the model is being developed, and this helps in creating the model boundaries of the system.

Trends: Once the agents, attributes, and environment are modeled, the next step is to look at trends that might already be taking effect. Most of the time, trends that are accounted for in the model are at a higher scale than the scale being modeled. This will be illustrated through the case study presented in Chapter 5. As an example, the socio-economic status (SES) of a particular school district in which one is intervening could be on the decline, or the student-teacher ratio in a particular school might be increasing due to political or economic changes. It is important to capture such trends and use them to model change in the system. This is demonstrated in the model design phase of the framework. While deciding which trends to capture, it is important to keep in mind the time horizon of the intervention. For those interventions which have a small time horizon (less than 2 years), some trends might be neglected for model simplification; but if the time horizon is large enough (3–5 years), then it might become important to incorporate trends in the attributes of the agents and the environment. This distinction of when to model trends is demonstrated through the two case studies presented in Chapter 3 and Chapter 5. The time horizon of the Chapter 3 case study was small enough such that modeling trends was not required, whereas the Chapter 5 case study is a longer intervention in which trends play a critical role in analyzing the implementation of that intervention. Also, the incorporation of trends guides the estimation of the weights in the change equations. The change equations are presented in the model design phase.

Acceptable zone: Another critical task in the model definition phase is to identify the criteria for intervention success. The 'Acceptable Zone' encapsulates the set of acceptable states of the system at the end of the intervention. To define this, it is important to work with the 'public policy researchers'. An acceptable zone can be constructed based upon the objectives of the intervention. Depending upon the purpose of the intervention, different levels of the agents' attributes and the environment would come under the umbrella of acceptable zone states. Defining the acceptable zone is very important as it directly affects our analysis of the barriers and enablers of implementing a successful intervention. Failure to identify the correct parameters for success of an intervention can shift the focus of the policy makers and practitioners toward a completely different set of attributes and parameters that might not actually be critical for understanding the success of the intervention. Success can also be defined via sustainability of the intervention. A sustainable intervention is one in which the intervention is carried forward in the school even after the intervening body, for example, the school partner implementing the intervention, leaves the school. Figures 8(a) and (b) illustrate state transitions from an initial state to an end state that fall within and outside the acceptable zone, respectively.



(a)



(b)

Figure 8: (a) Acceptable zone - Realized; (b) Acceptable zone - Not realized

For comprehensive interventions, defining the acceptable zone can be challenging. In this framework, the conceptual study presented in [30] is utilized to model the acceptable zone. There are three gaps affecting intervention implementation which are presented in this study: policy management, capability, and cultural gaps. Each of these three gaps is composed of a certain set of attributes and is at the system level. A gap is defined as the average of the difference between the actual state and the ideal state of the attributes it is composed of (a weighted average can also be taken). The policy management (PM) gap is the gap in the school attributes such as class duration, available supplies, teacher preparation time etc. The capability (Ca) gap is the gap in the attributes of the teachers/staff involved in the implementation of the intervention, such as teaching ability and content knowledge. The cultural (Cu) gap is the gap in the support for the intervention amongst the school agents. In each of these three dimensions there is an acceptable gap which constitutes the acceptable zone as shown in Figure 9.

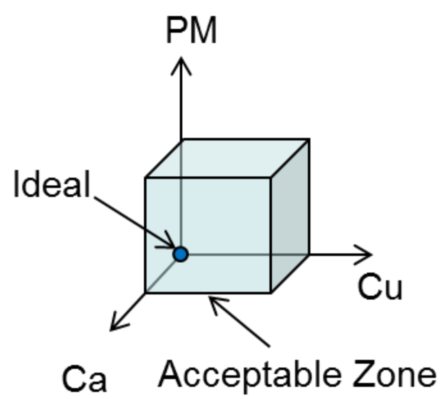


Figure 9: 3-dimensional acceptable zone using the gaps

Figure 9 describes the acceptable zone for a given system. Any system will have gaps in these three dimensions. In Figure 9, each axis represents one of the three gaps, and each gap has an acceptable tolerance. This tolerance can be calculated based upon the acceptable tolerance in the attributes that make up that gap, which is based upon the desired states of the attributes for a successful intervention. This tolerance is to accommodate the fact that any school system is not perfect, and there is always an acceptable gap which can still lead to a successful intervention. A required criterion for success is to have all three gaps within their acceptable tolerance. This criterion can also be defined on the basis of sustainability of the intervention. Another criterion to adhere to is the change in the probability of sustainability of the duration of the intervention. Ideally, the probability of sustainability should have improved by the end of the intervention, in comparison to what it was at the start of the intervention. The method to calculate probability of sustainability is presented in the section below.

Fitness function and probability of sustainability: It is useful to have a single parameter representing the overall fitness of the system in order to conduct the Method of Morris and sensitivity analysis. Based upon the above three gaps, an overall fitness function, to represent the health of the school in terms of sustaining this intervention, can be constructed as follows:

$$F = \frac{3}{PM + Ca + Cu} \tag{5}$$

where F represents the overall fitness. While the above fitness function is useful to see the overall change in the system, it may not be a good measure to compare different system performances due to the following observations:

1. Since the three gaps are added together in the above fitness function, individual gap effects get masked. For example: values of (0, 0, 2) and (0, 1, 1) for the three

gaps give the same fitness value. If each of the gaps above vary from (0, 2), then one could argue that the first system with gaps (0, 0, 2) should have a lower fitness value than the second system with gaps (0, 1, 1).

2. The absolute value of the overall fitness measure in Equation (5) does not mean much, as it varies from zero to infinity. A parameter is needed which can be related to the probability of sustainability of the intervention.
3. It does not incorporate trends in the three gaps.
4. The three gaps are weighted equally.

For these reasons, we develop a probability function taking the above criteria into account. We utilize a modified logit model which is used to predict binary response from continuous variables. The binary response here is the success (sustainability) or failure (lack of sustainability) of the intervention. This can be also considered a risk measure, where a higher probability of sustainability characterizes a lower-risk school environment in which to intervene and vice versa. The function used in the logit model is the logistic function:

$$l_0(x) = \frac{1}{1+e^{-x}} \tag{6}$$

where x represents the gap and $l_0(x)$ represents the logistic function value. Notice that $l_0(x)$ increases from $0 \to 1$ as x increases on the real line. We use a modified logistic function when gaps are the inputs to this function, i.e.,

$$l_1(x) = \frac{1}{1+e^{-(M-x)}} \tag{7}$$

where M is the maximum possible value of the gap. The modified logistic function $l_1(x)$ goes from $0.5 \rightarrow \frac{1}{1+e^{-M}}$ as x goes from $M \rightarrow 0$. So, $l_1(x)$ is decreasing in x, which is appropriate since the probability of sustainability should decrease as the gap increases. In order to convert $l_1(x)$ into a probability function, $p(x)$, the range of $l_1(x)$ can be shifted from (0.5, $\frac{1}{1+e^{-M}}$) to (0, 1) using the following transformation:

$$p(x) = \frac{l_1(x) - 0.5}{\left(1/_{(1+e^{-M})}\right) - 0.5} \tag{8}$$

Now, $p(x)$ goes from $0 \rightarrow 1$ as x goes from $M \rightarrow 0$. Furthermore, $l_1(x)$ can be modified to incorporate a trend in x:

$$l_2(x) = \frac{1}{1+e^{-(M-x-\frac{dx}{dt}\cdot\Delta t)}} \tag{9}$$

where $\frac{dx}{dt}$ is the change in x averaged across the previous time periods and Δt captures the time increment for which one wishes to incorporate trend effects. The quantity $l_2(x)$ can be converted into a probability function using the same transformation as described above. So the new $p(x)$ is:

$$p(x) = \frac{l_2(x) - 0.5}{\left(1/_{(1+e^{-M})}\right) - 0.5} \tag{10}$$

Another feature which is important to incorporate in the probability of sustainability is that larger gaps should be penalized more, especially if they are outside the acceptable zone gap tolerance. Let us assume that the gaps vary from (0, 2) and the acceptable tolerance is 0.67 (these values are taken from the case study presented in Chapter 5 and are explained in that chapter). So, to incorporate this feature, the probability function can be modified as follows:

$$p^*(x) = a(x) \cdot p(x) \tag{11}$$

where $a(x) = \begin{cases} 1; x \leq 0.67 \\ \frac{2-x}{2-0.67}; x > 0.67 \end{cases}$

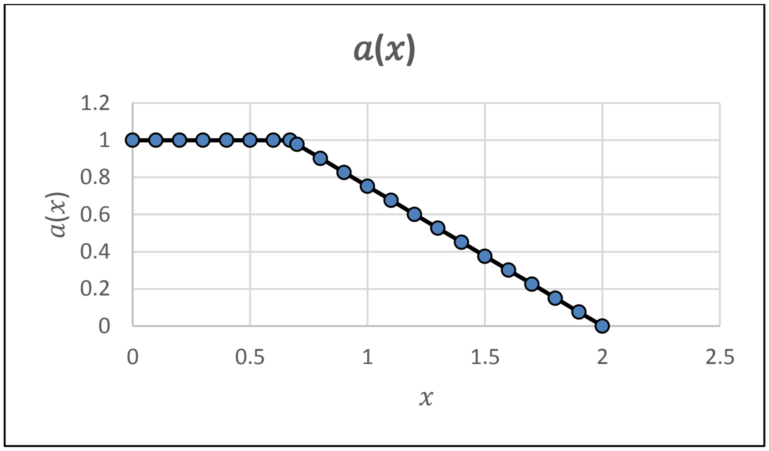


Figure 10: Graph of $a(x)$

From Figure 10, the effect of $a(x)$ on the probability function can be understood as follows: if x is beyond the acceptable threshold, then there is an additional penalty on the probability function, and this penalty increases as x increases. Here, $p^*(.)$ constitutes the probability of sustaining an intervention corresponding to a particular gap.

There are three gaps: policy management, capability, and cultural, and each of their individual probabilities have to be combined in order to calculate the probability of success of the intervention:

$$P(I) = P(PM, Ca, Cu) = w_{PM} \cdot p^*(PM) + w_{Ca} \cdot p^*(Ca) + w_{Cu} \cdot p^*(Cu) \tag{12}$$

where $P(I)$ is the probability of sustainability of the intervention, and w_{PM}, w_{Ca}, w_{Cu} are the weights corresponding to the three gaps; $w_{PM} + w_{Ca} + w_{Cu} = 1$. When a particular gap is large in magnitude, the corresponding probability function $p^*(.)$ will be low. For a successful intervention, all three gaps have to be low enough. Even if only one of the three gaps is large, it makes intuitive sense that the probability of sustainability of the intervention will be low. However, if the three weights are assigned equally to the three gaps, then even if one of the gaps is large (corresponding to $p^*(.)$ being low), then the overall probability of sustainability $P(I)$ might still be moderately high because of the other two gaps. One possible way to avoid this non-intuitive outcome is to use a weighting scheme similar to a harmonic weighted average. In a harmonic weighted average scheme, the weights are calculated as follows:

$$w_{PM} = \frac{\frac{1}{p^*(PM)}}{\frac{1}{p^*(PM)} + \frac{1}{p^*(Ca)} + \frac{1}{p^*(Cu)}}; w_{Ca} = \frac{\frac{1}{p^*(Ca)}}{\frac{1}{p^*(PM)} + \frac{1}{p^*(Ca)} + \frac{1}{p^*(Cu)}};$$

$$w_{Cu} = \frac{\frac{1}{p^*(Cu)}}{\frac{1}{p^*(PM)} + \frac{1}{p^*(Ca)} + \frac{1}{p^*(Cu)}}$$

$$and$$

$$P_H(I) = \frac{3}{\frac{1}{p^*(PM)} + \frac{1}{p^*(Ca)} + \frac{1}{p^*(Cu)}} \tag{13}$$

where $P_H(I)$ is the probability of sustainability of the intervention using harmonic weights. One drawback of taking the harmonic average is that, if any of the gaps are equal to 0 then $P_H(I)$ equals 0. There is no differentiation if there is one gap that is equal to 0 or two gaps or three gaps that are equal to 0. Another weighting scheme similar to the harmonic weighting scheme is proposed below to avoid this:

1. Assign the highest weight to the highest gap.
2. Assign the next highest weight to the next highest gap and so on.
3. Break ties randomly.

An example combination of weights that can be used is (0.7, 0.2, 0.1). Using the same range for gaps as before, i.e., (0, 2), and the acceptable gap threshold of 0.67 for each gap, the performance of the functions $P_H(I)$ and $P(I)$ is given by Table 10.

Table 10: Performance of $P(I)$ for various gap levels

Gap 1	**Gap 2**	**Gap 3**	$\boldsymbol{w_1}$	$\boldsymbol{w_2}$	$\boldsymbol{w_3}$	$\boldsymbol{P(I)}$	$\boldsymbol{P_H(I)}$
0	0	0	0.7	0.2	0.1	1.00	1.00
2	2	2	0.7	0.2	0.1	0.00	0.00
0.66	0.66	0.66	0.7	0.2	0.1	0.77	0.77
0	0	2	0.1	0.2	0.7	0.30	0.00
0	1	1	0.1	0.2	0.7	0.51	0.56
1	0.25	0.5	0.7	0.1	0.2	0.58	0.67

From the above table, we see that both $P_H(I)$ (with the harmonic weights) and $P(I)$ (with the combination of weights (0.7, 0.2, 0.1) and the dynamic weight assignment algorithm) satisfy the criteria listed above for measuring and comparing the performance of the system. Both of them avoid masking of the gaps when one of them is especially high and another is especially low; and the final probability value is intuitive to understand and easy to compare across different systems. Now, to compare $P_H(I)$ and $P(I)$, consider the following gap combinations: (2, 2, 2) and (0, 0, 2); $P_H(I)$ is 0 for both, whereas $P(I)$ is 0 and 0.30, respectively. Therefore, $P(I)$ might be preferred over $P_H(I)$ as it distinguishes between these two combinations.

From a policy analysis point of view, the above probability measures can also be used to quantify risk while intervening in a given school system. The school system can be categorized as a high-risk or low-risk school in which to intervene based upon the probability measure obtained above. A higher probability of sustainability will imply a low-risk environment and vice versa. As the intervention is being implemented, ideally the school will move from a high-risk to a low-risk environment. But this might not always be the case, and if the school moves from a low-risk to a high-risk environment, or even remains as high risk, then the particular intervention would not be considered sustainable. This can be another criterion to consider while defining the acceptable zone.

Another aspect to keep in mind, while using the above probability measure to quantify risk or probability of sustainability of the intervention, is the turnover rate amongst the school agents. If there is a high turnover rate, especially in an agent like the teacher which carries out the intervention, then the probability function should be further discounted when assessing the overall probability of sustainability of that intervention.

This concludes the model definition phase of the framework. Using the above steps, the model builder should be able to construct a network of the agents critical to model the intervention and identify the relationships and resource flows in the network. Now we move on to the next phase of the framework, which is model design.

4.3 Model Design

This is the second phase in the ESIM framework. In this phase, first the complete agent-based conceptual model of the intervention being modeled is developed. Then the conceptual model is validated with the help of subject-matter experts (SMEs) using the

4C's framework discussed in the model validation phase. After that, the computer simulation model is implemented and verified.

Agents' behaviors: To start building the conceptual model, it is critical to understand the behavior of each agent in the model. The behavior of an agent is different from the attribute of an agent. Attributes represent the characteristics of the agents whereas behavior represents the actions of the agents. It is not possible to capture all the behaviors of the agents, but it is important to model the appropriate behaviors which are induced by the intervention or which affect the intervention. Each intervention facilitates certain behaviors through interactions amongst agents and induces resource flows in the system. There are also behaviors of the agents taking place in the system which are not due to the intervention but which still affect the intervention implementation. A table should be developed containing the different types of agents and their roles or behaviors in the model. Table 11 is a simple illustrative example.

Table 11: Agents' behaviors

Agents	**Roles/behaviors**
School administration	• Creating annual school budget • Monitoring the principal's performance
Teachers	• Improving students' test scores • Interacting with the community • Interacting with the principal • Interacting with each other
Students	• Learning from the teacher • Interacting with each other • Feedback to the community

Table 11 is a small example of the types of roles different agents in the education system might play in school interventions. To develop such a table it is important to work with the "practitioners" group in the education system and others involved in the intervention. Once the behavior of the agents is clearly defined, rules can be developed to model change in the system.

Agent-based model rules: Rules govern the change in attributes of the agents and the environment. The framework being developed proposes modeling the change in the system's states as a Discrete-Time Markov Chain (DTMC). There are two main assumptions that must be made to model system changes as a DTMC. First, we assume that change can be modeled as taking place in discrete time steps. Second, we assume that changes in the agents' attributes depend only on the current attribute levels and current relationships and do not depend on past system states. This simplification helps in making the model tractable. To model this set-up, the time horizon for the intervention is divided into discrete time periods, during which the attributes of the agents and their relationships have some probability of change. There are three possible movements in the states: improving, no change, or worsening. The state change probability equation is made up of two components: an *internal* component to the intervention that corresponds to the phenomena captured in the model, and an *external* component that accounts for parameters external to intervention and outside the scope of the model which may affect the state change. Equation (14) represents the general structure of the changes taking place in the model:

$$p_{change} = w_{internal} \cdot p_{internal} \cdot f_s + w_{external} \cdot p_{external} \tag{14}$$

where p_{change} is the overall probability of change, $p_{internal}$ captures the 'modeled' aspects of change, and $p_{external}$ captures the aspects that are not modeled but still affect the change probability. These probabilities are vector quantities representing the three possible state changes, $[p_{improve}, p_{stay}, p_{worsen}]$, which add up to 1. In addition, there are weights, $w_{internal}$ and $w_{external}$, which quantify the percentage of change probability associated with each of the internal and external probability vectors. These are non-negative and the sum of these weights is always 1. Individual weights can be tuned for each model and their impacts can be investigated using sensitivity analysis, which will be discussed in subsequent sections. The initial values of these weights can be estimated based upon the scale of the intervention, and the impact of the external factors in the system. If the intervention is a small-scale intervention, such as the creation of an after-school club, $w_{internal}$ will be low. On the other hand, if the intervention is large-scale, involving complete school reform, $w_{internal}$ will be high (closer to 1). External factors and trends in the school system also affect the weights. If there are external pressures in a school, or its performance is declining over time, the change that can be achieved through the intervention in this system would be more difficult compared to a school system where such factors did not exist. This is demonstrated through the case study presented in Chapter 5, where there are external pressures in one of the school systems where the intervention is being applied. The $w_{external}$ in this case is higher ($w_{internal}$ lower) for the school system having such external trends, in comparison to the school system which did not have these external trends. Finally, the internal portion of the equation includes a multiplicative factor f_s which captures an 'S' curve pattern in learning. This curve represents the cumulative adoption of innovation or change in a complex adaptive system

[52]. The ‘S’ curve pattern models the behavior of a system where there is more inertia towards change when an attribute is at an especially low or an especially high level.

The internal component is further divided into two parts: *transient* and *steady state*. It is assumed that, before the intervention, the school is in a steady state. As implementation of the intervention begins, first the school goes into a temporary transient phase and then moves back to another steady state after a certain number of time steps. Equation (15) represents the structure of the internal component of the change equation:

$$p_{internal}(t) = e^{-kt}.p_{transient} + (1 - e^{-kt}).p_{steady} \tag{15}$$

where $p_{internal}(t)$ is the internal component of the change probability equation corresponding to time period *t*, k is the transient parameter affecting how long the system stays in the transient phase, and $p_{transient}$ and p_{steady} are the transient and steady-state parts of the internal component of the change probability equation, respectively. During the transient phase, there is a higher probability of change taking place in the system, as compared to the steady-state phase. The difference between these two components of the change equation is demonstrated in Chapter 5, when the specific change equations are presented for change in the attributes and relationships. As time becomes large enough, the internal component of the change probability equation becomes equal to the steady-state change probability:

$$\lim_{t \to \infty} p_{internal}(t) = p_{steady} \tag{16}$$

As we have found in this section, the above equations proposed to model change in the system are one way which has proven to be useful when applied to different case studies; however, the user of the framework is free to test different types of change equations while implementing the framework. An important avenue of future research is to test the framework using different change equations and analyze the effects on the model results.

Another parameter to determine while modeling change in the system is the step size. Step size is the level by which an attribute can change during a single time step either in the positive or negative direction. Step sizes can be chosen depending upon the time that an attribute needs to reach its maximum level. One possible way to determine step size is by setting the step size to max/T, where max is the maximum value of the attribute and T is the number of time steps required to change the attribute from zero to its maximum value. The time that it takes for an attribute to reach its maximum value is a subjective measure and should be decided in conjunction with the 'educational researchers'. Step size does not necessarily have to be the same for all attributes. As the number of time periods in which the time horizon is divided decreases, the length of a time step increases and so does the step size. Again, sensitivity analysis should be performed using different step sizes.

There are three main sub-models in the framework developed, all of which are inter-dependent in their dynamics of change. They are the changes in attributes, relationships, and resource flows. These are discussed next, along with the specifications of the DTMC.

Changes in attributes: The internal change probability, $p_{internal}$, for a particular agent is a function of three qualities: the current attributes of the agent, the agent's current relationships with other agents, and the attributes of the agents with whom the agent has relationships. If the agent has a relationship with the agency implementing the intervention, then the resource flows from the agency affect the change in attributes of this agent as well. Since different agents' attributes change over time, it is important to look at the intervention's agent network and also use the theories developed in education research to determine which agents' attributes affect the other agents' attributes. There are many studies in the literature that look at the different factors which affect student learning, teacher motivation, and school climate [53–59]. These studies analyze the effects of factors such as self-efficacy, grit, socio-economic status of the student population, teacher collaboration, responsiveness of administrators and many others. For inquiry-based interventions, [60] studies the effects of professional development on teaching practices and classroom culture. Such studies can be very helpful in guiding the user of this framework while building the inner mechanisms of the model which change the attributes of the agents. However, one should be careful in using the results of such studies, as some of them are context specific and the final *change in attributes* sub-model for the specific intervention should be developed working in parallel with the 'public policy researchers' and 'educational researchers' involved in the specific intervention.

Change in relationships: In addition to attributes, relationships also change over the course of the intervention: relationships can become more positive, stay the same, or may sour. Two concepts from social network theory aid in modeling this change in relationships: homophily and structural balance [61, 62]. Homophily assumes that individuals are more likely to form ties with other individuals who are similar to them in a variety of ways, including demographics, hobbies, and interests. Structural balance assumes that individuals are more likely to form positive ties with friends of friends, negative ties with friends of enemies, and positive ties with enemies of enemies. To model the teachers' social network, some of the other concepts that have been applied are proximity (physical distance), perception of expertise, and reform activities [25]. Proximity has been used to represent the physical distance between the work places of the teachers. Proximity has also been used to model the student social network, where proximity can represent the interaction opportunities the students get outside of the school, for example, if they live in the same neighborhood. Perception of expertise represents how often a teacher would approach another teacher to seek advice on a work-related problem. Reform activities represent the frequency of interactions between teachers as a result of the structure of the reform/intervention. Sometimes it even helps to divide the social network into two parts representing the *instrumental* (work-related) and *expressive* (non-work-related) relationships, or to just use one of these parts to model the relationships. Depending upon the context of the intervention, an appropriate social network structure should be chosen to model the relationships. It is important to construct a social network of agents consistent with the intervention as the attribute changes of the agents depend upon the relationships between the agents.

Resource flows: Resource flows are the final component modeled in the agent-based modeling framework. Resources could be information, money, or any other precious entity depending upon the school intervention being modeled. Within a school, these flows may take place through the relationships between the school agents and the intervening agency, and are impacted by the attributes of the agents and the strength of the relationships between the agents. An example of information flow could be the diffusion of an intervention amongst the teacher population and/or other agents. This process can be modeled similar to the disease-spreading models that have been extensively studied and applied to model epidemics in society. Another example could be the flow of money taking place amongst the agents. This is illustrated through the case studies in this dissertation. At a different scale, when the intervening agency is implementing the intervention across multiple schools, there can be varying allocations of resources to these school based upon the overall states of the schools, gauged using the three gaps: policy management, capability, and cultural, which we discussed during the development of the acceptable zone. Therefore, resources can be allocated based upon the gaps in the system. The resource allocation can also vary each year as the gaps in each school change. Since most of the interventions are constrained in their implementation by the resources available, it is important to model resource flows taking place in the system to better understand the barriers and enablers of the intervention.

Specifications of the Discrete-Time Markov Chain (DTMC): We use a DTMC to model change in the attributes and relationships of the agents. As mentioned before, there are two main assumptions that must be made to model system changes as a DTMC. First, it is assumed that change can be modeled as taking place in discrete time steps. Second, it is assumed that changes in the agents' attributes depend only on the current attribute level and current relationships and do not depend on past system states. This helps in keeping the model tractable, but at the same time, preserves the state-dependent property in the system, where the future states depend only upon the current state of the system. In order to completely define a DTMC, there are two main components that have to be determined: the initial state of the system, and the transition probabilities from one state to another. These are discussed below:

(i) The system state, at a particular time step, is a combination of the states of all the agents' attributes and their relationships (at that time step) being modeled for the particular intervention. To characterize the initial state of the DTMC at the start of the intervention (i.e., at time step t = 0), data are collected about the agents' attributes and relationships in the given school system where the intervention is taking place. This is used to populate the initial state of the DTMC and seed the model. For example, the level of inquiry teaching of a given teacher would be observed at the start of the intervention. This would be the initial state of that attribute of that agent in the model.

(ii) The next step is to characterize the transition probabilities in the system. At any given time step, an agent's attributes and the relationships it has with the other agents in the system can all have three possible state changes: improve, stay the same, or worsen. Each of these state changes has a probability associated with it $[p_{improve}, p_{stay}, p_{worsen}]$, which add up to 1. These probabilities depend upon three things: the agent's own attributes, the relationships it has with the other agents, and the other agents' attributes.

First, let us look at the change in the attributes of the agents. If the attribute level improves or worsens in a given time step, it is by a given step size, which was discussed earlier in this section. Let this step size be denoted by Δa. Let the attribute be at a level a_t at time period t and the range of values that the attribute can take be between $[0, M]$, where M represents the maximum level of the attribute. Then the transition probabilities in the state of this attribute are given as follows:

$$p_{ij} = P\{a_{t+1} = j | a_t = i\} = \begin{cases} p_{improve}; if\ j = a_t + \Delta a \\ p_{stay} \quad ; if\ j = a_t \\ p_{worsen}\ ; if\ j = a_t - \Delta a \\ \quad 0 \qquad ; else \end{cases} \tag{17}$$

Here p_{ij} represents the state transition probability of the attribute at any given time step, between the states i and j. As seen from Equation (17), between consecutive time steps, the attribute can either improve by the step size Δa (with probability $p_{improve}$), or stay at the same level (with probability p_{stay}), or decrease by the step size Δa (with probability p_{worsen}). All the other state transition probabilities are zero. The state transition for the particular attribute can also be understood via Figure 11.

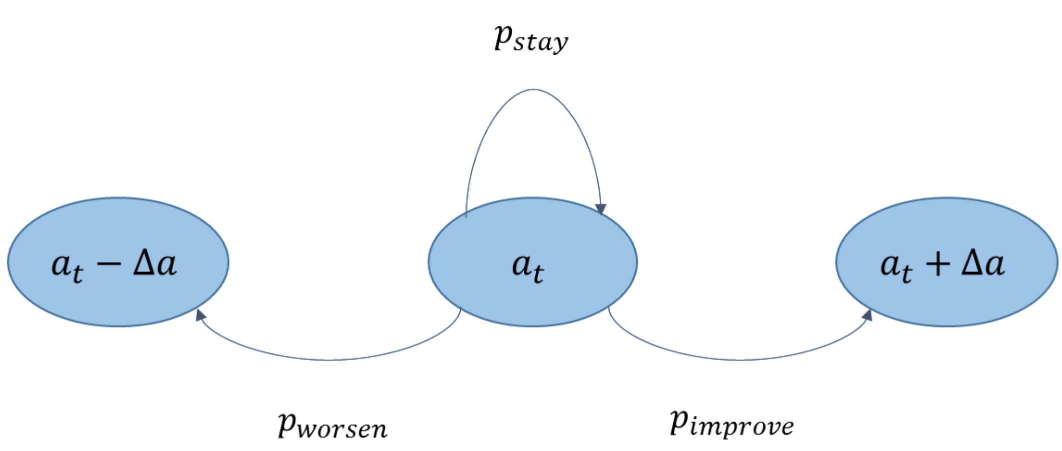


Figure 11: Attribute state transition between consecutive time steps

The state transition probabilities are also required to satisfy the condition:

$$\sum_{j \in [0,M]} p_{ij} = 1 \tag{18}$$

The above condition simply says that the attribute will transition *somewhere* from state i (including the possibility that the attribute stays at state i).

Similar to the above method of modeling change in the attributes of the agents, the change in the relationships amongst the agents can also be modeled. Let the change in a particular relationship take place with a step size Δr. Let the relationship be at a level r_t at time period t and the range of values that the attribute can take be between $[0, R]$, where R represents the maximum level of the relationship. Then the transition probabilities in the state of this relationship are given as follows:

$$p_{ij} = P\{r_{t+1} = j | r_t = i\} = \begin{cases} p_{improve}; if\ j = r_t + \Delta r \\ p_{stay} \quad ; if\ j = r_t \\ p_{worsen}\ ; if\ j = r_t - \Delta r \\ \quad 0 \qquad ; else \end{cases} \tag{19}$$

Here p_{ij} represents the state transition probability of the relationship, at a given time step, between the states i and j. As seen from Equation (19), between consecutive time steps, the relationship can either improve by the step size Δr (with probability $p_{improve}$), or stay at the same level (with probability p_{stay}), or decrease by the step size Δr (with probability p_{worsen}). All the other state transition probabilities are zero. The state transition for the particular relationship can also be understood via Figure 12.

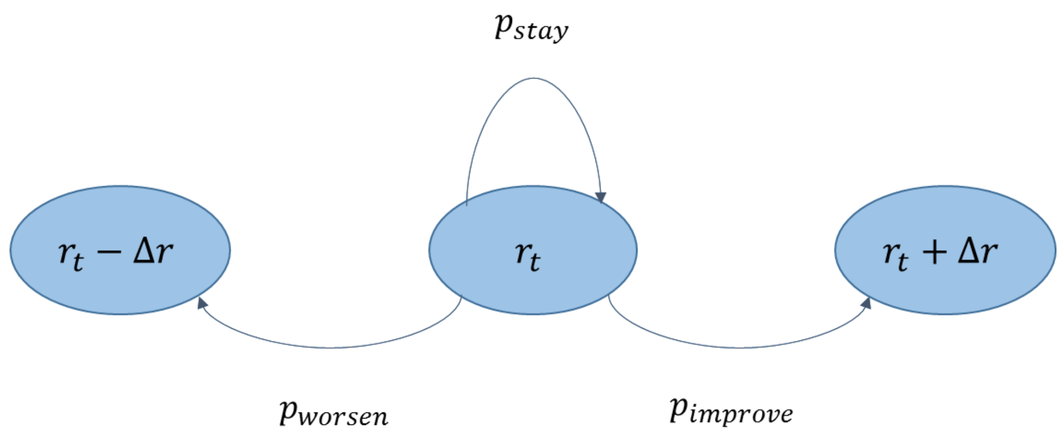


Figure 12: Relationship state transition between consecutive time steps

The above figure demonstrates that, between any two consecutive time steps, the relationship can either stay at the same level, improve by the step size, or decrease by the step size. Again, the state transition probabilities are also required to satisfy the condition:

$$\sum_{j \in [0,R]} p_{ij} = 1 \tag{20}$$

The above condition simply says that the relationship will transition *somewhere* from state i (including the possibility that the relationship stays at state i).

The path of the Markov Chain can be simulated using the above conditions, starting with the initial state of the system, and then mapping the change in the system state at each time step using the transition probabilities for each attribute and relationship constituting the system state. The three change probabilities: $p_{improve}$, p_{stay}, and p_{worsen} are calculated based upon the internal and external change probability components discussed before in this section. This completes the definition of the DTMC used to model the change in the system.

<u>Conceptual model validation</u>: Once the agents' behaviors are documented and the three sub-models (*change in attributes, change in relationships,* and *resource flows*) are constructed, the conceptual model is complete. The conceptual model contains all the information about the agents, agents' attributes, environment, and how the system evolves over time.

All the tasks completed in the model definition phase until now were concerned with building the conceptual model. It is very important before moving forward to get the conceptual model validated with the assistance of SMEs who are the 'practitioners' and 'educational researchers'. This is done using the 4C's framework described in the model validation phase of the framework below. The *four C's* stand for Completeness, Consistency, Coherence, and Correctness. Each of these is defined later. During this step, the model builder might have to go back to the model definition phase or the previous steps discussed in the model design phase and refine certain attributes of the agents being modeled or make changes in the agent-based model rules. It is more efficient to make these changes now before the implementation of the simulation model for analysis. Hence, the interconnectedness between the different phases of the framework while building the model arises at this stage.

<u>Data collection</u>: In education system models it is critical to carry out the data collection step early so that one does not make any assumptions in the model about the availability of data without first verifying that issue. Building the conceptual model by working in parallel with ‘educational researchers’, ‘practitioners’, and ‘public policy researchers’ ensures that the model reflects the reality of the education system, and that data exist to populate the initial conditions of the model. Still, when the data are collected, there could be missing or uncertain information for certain parts of the model. This would either result in refining or modifying the conceptual model or putting a red flag on those input variables for which there is a excessive of uncertainty in the data. Such variables should later be put to a sensitivity analysis test during the model analysis phase to estimate the effects of uncertainty in their values on the model’s output. It always helps to analyze those types of school interventions for which there is much readily available data about the agents’ attributes and relationships; but such an ideal scenario rarely exists in reality and data collection is something that every model builder will have to face while applying the framework to model the education system.

Computer simulation model: After the construction of the conceptual model is complete and data are collected for seeding the model, the next step is to develop the simulation model. Modeling software packages exist such as NetLogo, Repast, and InsightMaker etc. to program agent-based simulation models. Since object-oriented programming is particularly suited for developing agent-based models, programming languages like C# or Java could also be used to develop such models. Both the pre-existing software and object-oriented programming languages have their advantages and disadvantages. The advantages of agent-based modeling software are that the models are quick to implement and they have many different analysis tools to analyze the model results. However, since they are pre-developed, there is a limitation in terms of the various features of the model that can be built with them. The object-oriented programming languages on the other hand provide complete flexibility to the model builder but demand experience from the modeler in using such languages to build an agent-based model.

Once the computer simulation model is built, the next step is to verify the model for correctness. There are two main conditions that the simulation model has to satisfy for verification: there should be no programming bugs in the model and it should have the desired behavior which the conceptual model intended. These steps are discussed in the model validation phase. Also, the model analysis phase is very good at catching programming errors in the simulation model developed, as sensitivity analysis and extreme condition tests verify that the simulation model produces outputs in coherence with the conceptual model's intent. Once the simulation model is ready for use, the next phase of the framework, the model analysis phase, is entered.

4.4 Model Analysis

In this phase of the framework, simulation results are generated and then analyzed using various techniques. While simulating outcomes that are consistent with reality is helpful, the real contribution of this framework is two-fold: the sensitivity analysis of the model and the determination of factors that most greatly affect the outcome/intervention. These are discussed in detail below.

Simulation results: To generate simulation results, the initial state of the computational model should be populated using the data collected for the intervention. Since the agent-based model is built using a Discrete Time Markov Chain, it is a stochastic model, so multiple simulation runs must be made to analyze the output. The variance in the model results should be reported for the number of runs made. While conducting sensitivity analyses and other types of tests on the model, it is important to keep in mind that this is a stochastic model and the changes in the model output can be skewed because of the model's natural variance. The technique proposed in the framework to overcome this challenge is to use the average of multiple independent simulation runs as one data point. Assuming *n* simulation runs correspond to a single data point, variance estimates should be supplied in the model results. The value of *n* should be chosen such that this variance is sufficiently low in order to be able to treat the model as reasonably deterministic. In this way, the change in the output can be attributed to the change in the inputs with a sufficient level of confidence, so that subsequent analysis of the model results can be performed.

Sensitivity analysis: To assess the model dependency on uncertain input variables and parameters such as weights and the S-curve parameter, a sensitivity analysis of the results with respect to the model parameters should be conducted. This is important to determine three main characteristics of the model results:

1. Sensitivity of the model results with respect to changes in parameter values
2. Reliability of the model results with respect to the uncertainty in parameter values
3. Robustness of the model results with respect to changes in model structure

Sensitivity refers to the change in model results with respect to the change in inputs, reliability refers to the confidence in model results under uncertainty, and robustness refers to the dependence of model results on the essentials of the model and assumptions made [63, 64]. Such analyses also help in the verification and validation of the model as discussed in the model validation phase.

<u>Method of Morris</u>: In this second phase of analysis, the most-significant attributes, or model inputs, are identified for the case being modeled. The analysis is undertaken using the Method of Morris (MoM) [47]. This method is useful for determining a subset of input variables, from amongst a larger set, which most likely have a significant impact on a particular outcome. The experimental plans in MoM are composed of individually randomized one-factor-at-a-time designs in the input variables. Conigliaro *et al.* provide a discussion about the advantages of MoM over other factorial sampling techniques and a summary of the method which is presented below [65]. One shortcoming of the MoM is that it does not provide output variance decomposition. This is acceptable given the purpose of this analysis in the framework, since MoM is being used simply to identify the inputs that have a statistically significant impact on the model outputs of interest.

MoM examines the changes in an output based upon experimental plans composed of randomized designs of the input variables. In each run only one input variable is given a new value allowing change in the output to be unambiguously attributed to change in that input. The mean and variance of the elementary effect of the input variables on the output is estimated. One sample of the elementary effect for the i^{th} input variable is defined as:

$$d_i(x) = \frac{[y(x_1,\dots,x_{i-1},x_i+\Delta,\dots,x_k)-y(x)]}{\Delta} \tag{21}$$

where x is a k-dimensional vector of model inputs, y is the model output being analyzed, $d_i(x)$ is the elementary effect for the i^{th} input variable, and Δ is often chosen as:

$$\Delta = \frac{p}{2\cdot(p-1)} \tag{22}$$

where p is the number of grid levels in the region of experimentation. The discrete random variable, of elementary effects associated with the i^{th} input variable, obtained by randomly sampling different x's, is denoted by F_i with mean μ_i, and standard deviation, σ_i. Since x produces a simple random sample, from each F_i, the mean and standard deviation of the observed elementary effects for input i are unbiased estimators of the mean and standard deviation of F_i, and the standard error of the mean can be estimated as $SEM_i = \sigma_i/\sqrt{r}$; where r is the number of random values for each input. Input factors with a large mean are likely to have an overall important influence on the output, while input factors with a large standard deviation may have interactions with other factors or may have non-linear effects.

To create a MoM design experiment, sampling matrices denoted B^* are constructed using the following equation:

$$B^* = \left(J_{k+1,k} \cdot x^* + \frac{\Delta}{2}\left[\left(2 \cdot B - J_{k+1,k}\right) \cdot D^* + J_{k+1,k}\right]\right) \cdot P^* \tag{23}$$

where $J_{k+1,k}$ is a $k+1$ by k matrix of ones, x^* is a randomly chosen base value of x, D^* is a k-dimensional diagonal matrix in which each element is either 1 or -1 with equal probability, and P^* is a k by k random permutation matrix. One sampling matrix, B^*, is needed for every sample of a main effect. If one desires n main effect samples, then $n*(k-1)$ function evaluations are needed.

This provides an overview of implementation of the MoM experiment as presented by Conigliaro *et al.* [65]; for detailed discussion about each step see the original paper by Morris [47]. MoM is a very useful technique and can be applied in determining the main attributes and relationships that affect the school intervention being modeled. MoM can be conducted on various output variables to analyze different objectives during the intervention. MoM can also be used to reduce the complexity of the model to perform validation on a reduced parameter and input space.

This concludes the model analysis phase of the framework. As this phase is implemented and the results analyzed, insights will be gained about the barriers and enablers of the intervention. This might also force the user of the framework to go back to the previous phases to refine or modify the conceptual and computational model. Now, model validation, the final phase of the framework, is presented.

4.5 Model Validation

This is the last phase of the framework and is very important in developing confidence about the model being built amongst its users. Even though the model validation phase is presented last, as mentioned previously, it goes hand in hand with the other phases of the framework. Waiting until the end to start the verification and validation process can lead to too many modifications in the model which could be expensive to implement at this stage; hence it is important to follow the trajectory shown in Figure 5 while applying the framework. The verification and validation steps implemented during the different phases of the framework are discussed now.

The 4C’s model validation: Validation of the conceptual model should be done with the aid of subject-matter experts who are the ‘practitioners’ and ‘educational researchers’, under the *4C’s* framework proposed by Pace for developing simulation conceptual models [66]. The *four C’s* stand for Completeness, Consistency, Coherence, and Correctness. The questions that the SMEs should ask while establishing the validity of the conceptual model for each of the *four C’s* are discussed below:

1. Completeness: Does the conceptual model contain all of the entities such as the agents, attributes, and their relationships essential to model the particular school intervention? Does the conceptual model capture all the processes and mechanisms happening during the implementation of the intervention?
2. Consistency: Are the entities and processes within the conceptual model compatible with the reality? Are the entities and processes within the conceptual model consistent with the scale at which the model is being built?
3. Coherence: Do all the agents, attributes, and relationships modeled have a function in the simulation (i.e., are there no extraneous factors being modeled)? Do all the elements modeled have potential (i.e., are there no parts of the conceptual model which are impossible to activate)?
4. Correctness: Is the conceptual model appropriate for modeling the particular school intervention? Does the conceptual model have the potential to fully satisfy the simulation requirements?

Answering the above questions about the conceptual model helps in refining the model and may result in a number of useful changes to the model. Developing a conceptual model which is fully complete, consistent, coherent, and correct is completely not possible in reality, but undertaking the above validation exercise for the conceptual model helps in building a more-useful and useable model. Now, a number of verification and validation techniques applicable for validating the education system models are discussed. These were compiled by Sargent in his work on validation and verification of simulation models [67].

Animation/graphics: The model's operational behavior is displayed graphically as the model moves through time. While modeling the education system, this could mean graphing the model's variables of interest over time.

Comparison to other models: Various results (e.g., outputs) of the simulation model being validated are compared to results of other (valid) models. This can currently be challenging to do for simulation models of the education system, since there are very few such models that exist. However, as more models are developed to study the education system and this framework is applied across various settings, this will be a much more achievable task.

Degenerate tests: The degeneracy of the model's behavior is tested by appropriate selection of values of the input and internal parameters. For example, the effect of teachers' content knowledge on students test scores could be analyzed by increasing the former variable. Degenerate tests are those which should have obvious results, but are conducted none-the-less to validate the model.

Event validation: The "events" of occurrences of the simulation model are compared to those of the real system. For example, compare the occurrences of events during a school intervention which is constrained in its occurrences by the resources available in the system. An example of this is illustrated through the case study modeled in the next chapter.

<u>Extreme condition tests</u>: The model structure and outputs should be plausible for any extreme and unlikely combination of levels of factors in the system. This goes back to the sensitivity analysis tests in the model analysis section, where the effect of a range of different parameters values is analyzed on the model outputs.

<u>Face validation</u>: Asking individuals knowledgeable about the system whether the model and/or its behavior are reasonable. This again makes collaboration with the ‘public policy researchers’, ‘educational researchers’, and ‘practitioners’ necessary and valuable for building a valid and useful model. The conceptual model validation using the *4C’s* framework falls under the category of face validation.

<u>Historical data validation</u>: If historical data exist (or if data are collected on a system for building a model), part of the data can be used to build the model and the remaining data are used to determine (test) whether the model behaves as the system does. For a school intervention, data collected throughout the implementation of the intervention should be used to compare the simulation results with the actual system behavior.

<u>Internal validation</u>: Several replications (runs) of a stochastic model are made to determine the amount of (internal) stochastic variability in the model. This goes back to the simulation results section of the model analysis phase where it is suggested that several runs of the model be made to determine the internal variance in the model.

Multistage validation: Naylor *et al.* proposed combining three historical methods of rationalism, empiricism, and positive economics into a multistage process of validation [68]. This validation method consists of (1) developing the model's assumptions on theory, observations, and general knowledge, (2) validating the model's assumptions where possible by empirically testing them, and (3) comparing (testing) the input-output relationships of the model to the real system.

Parameter variability – Sensitivity analysis: This technique consists of changing the values of the input and internal parameters of a model to determine the effects upon the model's behavior or output. This is discussed in detail during the model analysis phase of the framework.

A combination of the above techniques should be used to verify and validate the model built. Data validation can also be performed on a smaller set of model inputs and parameters after discerning the most-important attributes and relationships using the Method of Morris. This completes the development of the framework to model education systems. Application of the ESIM framework to a case study is presented in the next chapter. Applying the framework to model different school interventions helps in analyzing the barriers and enablers for successful interventions as well as improving the framework to make it more comprehensive.

CHAPTER 5

APPLICATION OF ESIM TO SLIDER

In the previous chapter, the ESIM framework to model interventions in the education system was presented. The framework has four different phases: model definition, model design, model analysis, and model validation; each of which was discussed in detail.

In this chapter, the ESIM framework is applied to a case study of an inquiry-based physical science intervention. First an introduction to the case study is given along with reasons to select this intervention, and then the application of the different phases of the framework to develop a model for this case study are discussed.

5.1 Introduction to SLIDER

SLIDER stands for Science Learning: Integrating Design, Engineering and Robotics. The following excerpt from the SLIDER NSF proposal document [69] gives a brief overview of this intervention:

"The Science Learning: Integrating Design, Engineering and Robotics (SLIDER) program is a collaboration between K-12 educators, university faculty, and educational outreach specialists located at Georgia Tech's Center for Education Integrating Science, Mathematics and Computing (CEISMC); learning theory and cognitive science professionals in Georgia Tech's Center for the Enhancement of Teaching and Learning (CETL), School of Psychology, School of Biomedical Engineering, and College of Computing; the Georgia Department of Education; and three Georgia school systems- one urban, one rural, and one suburban. SLIDER proposes to design and implement an

8th-grade Physical Science inquiry based curriculum that uses engineering design and LEGO Robotics as the context with which to teach science content and process skills. We will thoroughly research its effect on student learning, motivation, creativity, and problem-solving skills, and on longitudinal student academic achievement."

SLIDER is a five year intervention which started in 2010. The intervention was planned to be implemented across three demographically varying schools. Due to IRB constraints, the names of the schools cannot be disclosed in a publication, therefore, , so the following nomenclature will be used to represent the three schools:

- **School 1:** Rural school district school
- **School 2:** Suburban school district school
- **School 3:** Control school in sub-urban school district

This is an interesting case study to model because of the resulting varying levels of success of implementation across the three schools. School 3 was treated as a control variable since the agents' attributes and the environment attributes were already at or close to the highest levels, and stayed at that level throughout the intervention. On the other hand, the other two schools had a lot of variation in the attributes of its agents and the environment during the course of the intervention. The implementation of the intervention is modeled individually at these two schools and then at the end, all three of the schools are modeled together to analyze the proportion of resource allocations across the three schools.

5.2 Application of ESIM

The framework was applied to the SLIDER case study in collaboration with the team of educational researchers, public policy researchers, and industrial and systems engineering researchers at Georgia Tech. The modeling cycle followed was as shown in Figure 5, however application of each phase of the framework is discussed here sequentially starting with model definition. There are some differences in the number of teachers and students across School 1 and School 2, but apart from that, the type of agents, attributes, relationships, resource flows, and change equations are the same. Therefore, for the most part, the application of the framework to these two school is presented together, but some parts like trends, results are presented separately.

5.2.1 Model Definition

Problem definition: This intervention involved the design and implementation of an 8th grade physical science inquiry-based curriculum. During the first two years, capacity building in the schools was the focus along with the design of the curriculum, and in the remaining years, the focus was on implementing the curriculum across the schools and making changes in the design accordingly. Professional development was being imparted to the schools throughout the intervention to improve teachers' attributes so that the intervention could be sustained after Georgia Tech leaves the schools. The goal is to analyze the change in the state of the schools during the intervention and understanding what factors led to a successful implementation in one case and an unsuccessful implementation in the other case.

Scale of analysis: The scale of analysis of this intervention is at the grade level, when we are looking at the schools individually. All the 8^{th} grade sections at School 1 and School 2 were involved in this intervention. This is not a hierarchical intervention where one agent controls the dynamics of the majority of the system behaviors. The hierarchical structure breaks down to a structure where any two agents who are linked together can have resource flows and affect each other's attributes.

However, at the end of this chapter, when the schools are modeled together, the scale of analysis is one level higher. Each school is considered as a single agent. When analyzing the system across different scales, the micro-behavior at a lower scale gives rise to macro-behavior at a higher scale. This is demonstrated in *Section 5.3* of this chapter.

Temporal scale: The time horizon of this intervention is five years, with the focus on curriculum design in the first two years and on classroom implementation in the next three years. The time step for the first two years is taken as one month, but once the classroom implementation begins, there are two time steps being used. During the first three months (or twelve weeks) of the school year, the time step is taken as one week because the curriculum is being implemented during this time and the frequency of interaction amongst the agents increases. During the remaining nine months of the school year, the time step is taken as one month. This is another flexible feature of the framework, where the time step does not have to be fixed throughout the intervention. The above time steps were decided in consultation with the 'educational researchers' to allow sufficient time for an agents' attribute to change.

Agents, attributes and environment: First, let us look at the network of agents at different scales of analysis in Figure 13.

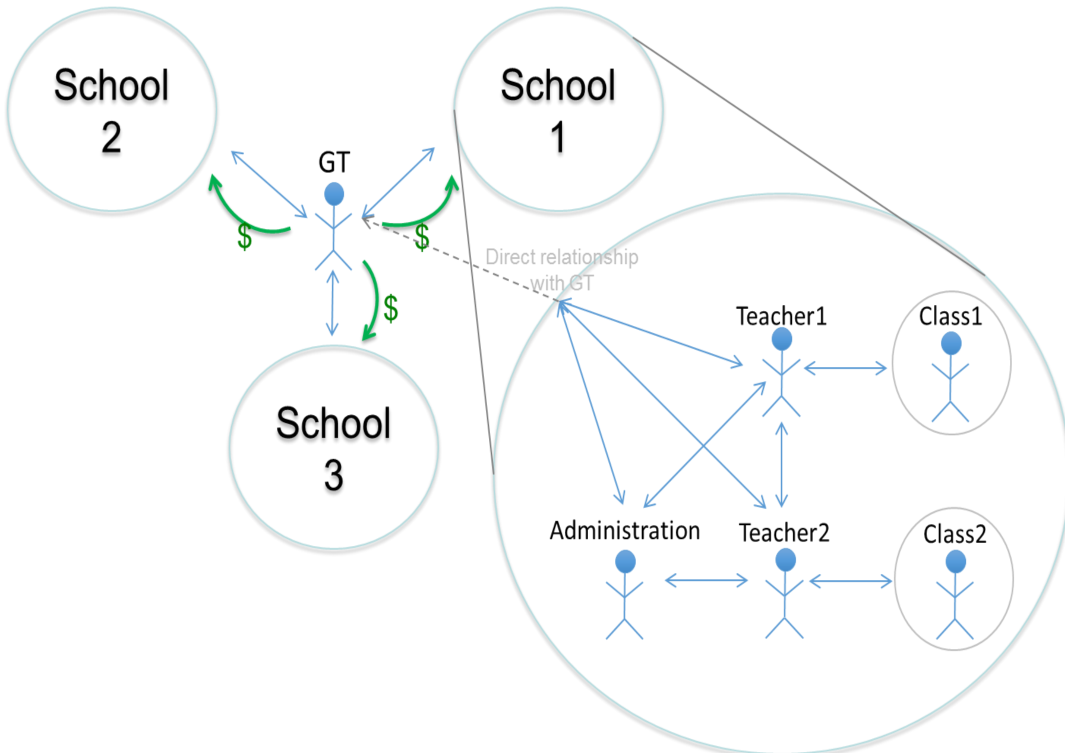


Figure 13: Agent network for SLIDER case study

Georgia Tech (GT), which is the agent implementing the intervention, is interacting with the three schools. The agent network within each school is shown. There are teachers who are implementing the intervention, administration which has relationships with both Georgia Tech and the teachers in the school. One thing to note is that, the classroom implementation starts from year 3 and also that all the class sections that a particular teacher teaches is lumped into one big class. Since there is very little deviation between the attributes of the different sections which a particular teacher teaches, the students across these sections are put together for modeling simplification. Also, since this class is modeled as a single agent, we take an average over the students in the class to get the class level attributes. The number of students that each teacher taught across different sections is around 100. The blue arrows in the above agent network represents a relationship that is being modeled in the case study, and the green arrows represent the resource flows. The agents being used in the above network were decided in consultation with the ‘educational researchers’ and ‘practitioners’ of the intervention.

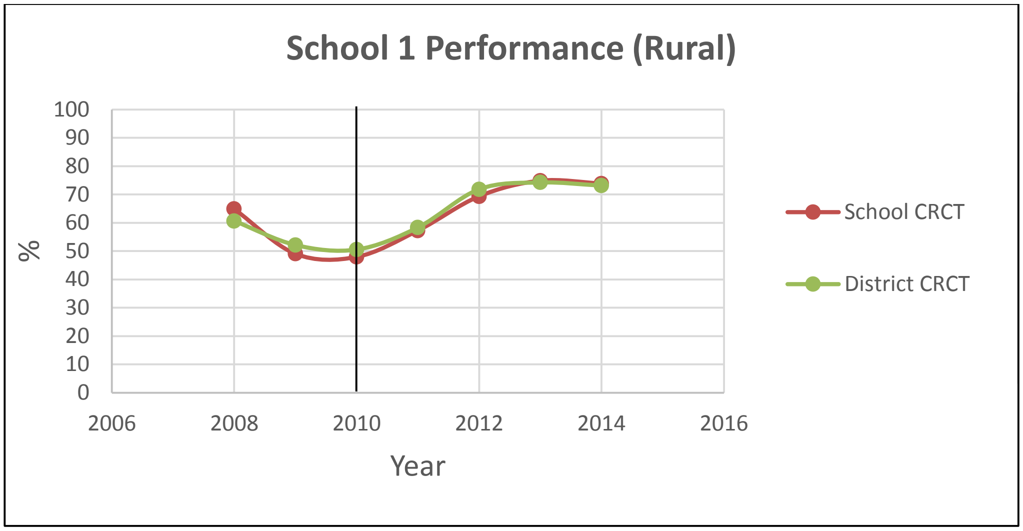
Now let us look at the attributes of the different agent classes and the environment (School), being modeled for the SLIDER case study:

- **Class:** SLIDER pre-post test scores (for tests given within the SLIDER curriculum), % free and reduced lunches, number of students
- **Teacher:** Inquiry teaching level, content knowledge, self-efficacy, organization citizenship, support for intervention
- **Administration:** Leadership, support for intervention
- **School:** Class duration, classroom space, student-teacher ratio, teacher preparation time, availability of supplies, teacher turnover rate, administration turnover rate, relative performance to other schools in the same district
- **Georgia Tech:** Total budget, professional development budget allocation, material supply budget allocation

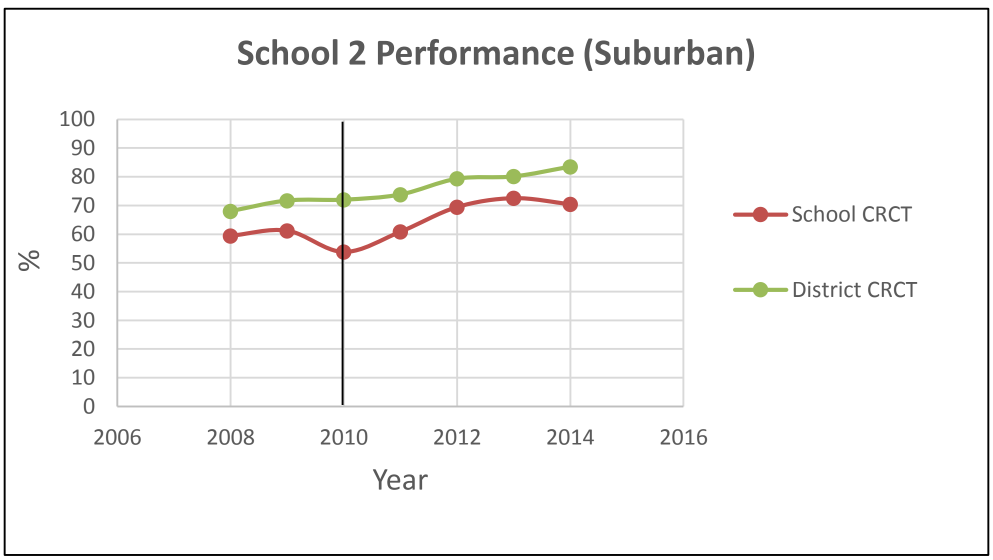
The above list of attributes is by no means comprehensive, but these are the preliminary set of attributes that were considered important to be modeled specific to this case study. These were decided based upon the inputs and discussions with the ‘educational researchers’ and ‘practitioners’. One of the attributes of the school is the turnover rate, in both the teachers and administration. The assumption to model turnover rate is that when the turnover happens, the agents which replace the current agents have attribute levels close to the current agents at the time of the start of the intervention. Also, turnover is modeled as a random process, where there is a certain probability of the agent leaving the school system at the end of each academic year. This probability is characterized by the turnover rate. After the Method of Morris experiment is completed, we get a smaller subset of attributes to consider in more detail.

The environment in this case study is made up of the school and the school district in which the intervention is being implemented. It also contains the relationships amongst the agents and the resource flow arcs between the agents. As mentioned before, in Figure 13 the blue arrows represent the relationships and the green arrows represent resource flows. There are some agents that are not connected, that does not mean that there might not be any interaction amongst them, but for the purposes of modeling this intervention, those relationships were not considered strong enough to change the outcomes of this case study. The resource flow arcs in the agent network represent the investment from Georgia Tech in terms of professional development and materials supply. These resource flows impact the attributes of the school agents and environment.

Trends: External trends were considered for School 1 and School 2, in the student-teacher ratio, and 8^{th} grade Science CRCT scores. The test scores levels at the two schools were also compared with the overall district performance to assess external pressures on the schools.



(a)

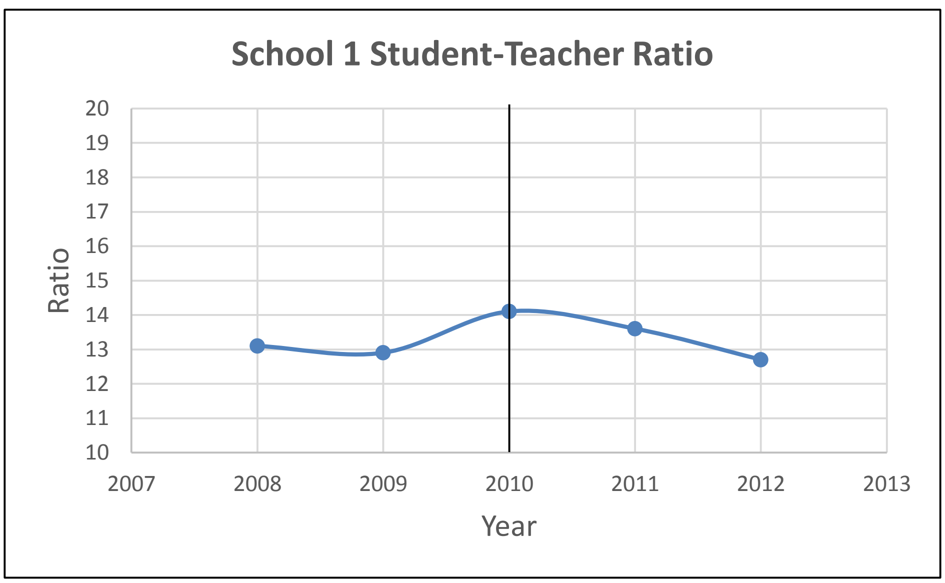


(b)

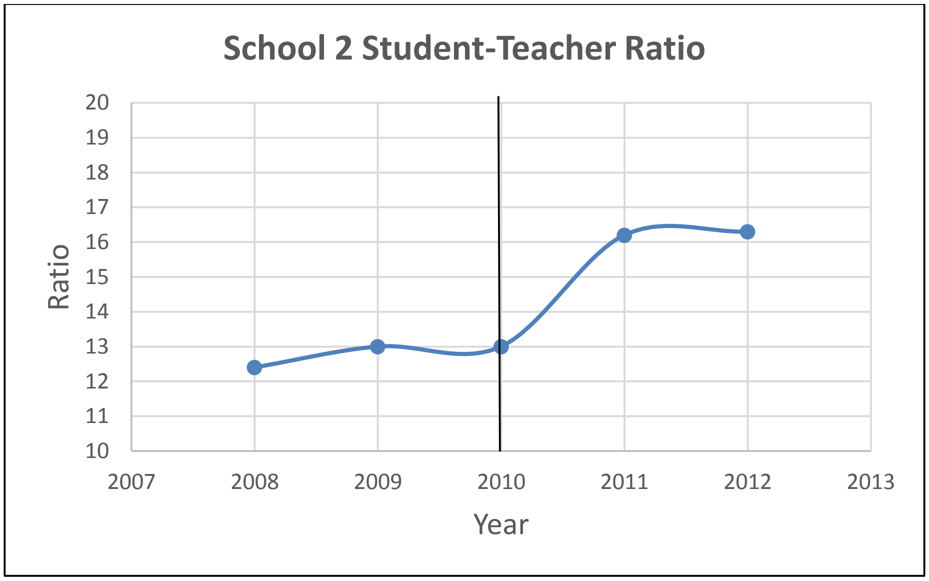
Figure 14: (a) School 1 8th grade CRCT scores; (b) School 2 8th grade CRCT scores

In above graphs in Figure 14, the key thing to note is the difference between the school and the district performance. For School 1 this difference is negligible, since there is only one other school in the district and their performance is similar. However, for

School 2 the difference is substantial and it was at its largest when the intervention started. There is a concept in organizational behavior called 'threat rigidity' [70] which can be applied in the case of School 2 because of the gap between the school and the district performance. Threat rigidity explains that whenever there are external pressures on an organization, the organization stops innovating and responding to new changes and reverts back to the tried and tested norms. This can guide the assessment of the *internal* and *external* weights in the change equations. From Figure 14 it can be inferred that there are higher external pressures on School 2 because of its lagging performance compared to the district and hence, the external weight for School 2 should be larger than of the one for School 1. The other trend observed was in the student to teacher ratio, which is shown below. This is the ratio of the total number of full-time teachers to the total number of students in the school.



(a)



(b)

Figure 15: (a) School 1 student-teacher ratio; (b) School 2 student-teacher ratio

It can be seen from Figure 15 that the student to teacher ratio at School 1 is pretty constant with some minor variations, whereas there is a huge spike in the student-teacher ratio at School 2, after the start of the intervention. This was when the economy had collapsed in 2010 and as a result, there were many teacher lay-offs in School 2's district. The above trends help guide the model analysis and results without having to account for them internally in the model.

Acceptable zone: As described in Chapter 4, the three gaps: policy management, capability, and cultural gaps can be used to quantify a three dimensional acceptable zone. The attributes in the case study that make up these gaps are the following:

- **Policy management gap:** Class duration, average class size (student-teacher ratio), classroom space, teacher preparation time, availability of supplies

- **Capability gap:** Teacher's inquiry teaching skill, teacher's content knowledge
- **Cultural gap:** Administration support for intervention, teacher support for intervention

The above gaps are calculated at the end of each year to assess the performance of the schools in comparison to the acceptable zone. The gaps in the different attributes which make up each of the gaps are weighted equally while calculating that particular gap. For each attribute, the acceptable tolerance is taken as the difference between the highest and the second highest level. Since the attributes are measured from (0, 2) with a 4-level scale, the acceptable tolerance is taken as 0.67. The change in the gaps across different years of the intervention and the final state guide whether or not the intervention can be sustained in that school.

As discussed in Chapter 4, the three gaps are also used in the calculation of the overall fitness function and the probability of sustainability function. The overall fitness function is used to conduct the sensitivity analysis and the Method of Morris experiment. The probability of sustainability function on the other hand is used to quantify risk of intervening in a particular school system and assess the chances of the intervention being sustained in that school.

5.2.2 Model Design

<u>Agents' behaviors</u>: Each agent modeled in the SLIDER case study has a specific role in the intervention. The administrators are the agents who facilitate change in the school and provide support to the teachers, the teachers are involved in the implementation of the intervention, Georgia Tech is designing the curriculum and providing resources, etc..

Table 12 summarizes the behaviors being modeled for the different agent classes based upon their roles in this case study.

Table 12: Agents' behaviors for the SLIDER case study

Agent	Behaviors
Class of students	• Participate in the intervention • Learn the curriculum being taught
Teacher	• Participate in professional development being given by Georgia Tech • Implement the curriculum in their classes • Participate in a relationship with Georgia Tech
Administration	• Create an environment for change in the school • Provide support to the teachers • Participate in a relationship with Georgia Tech
Georgia Tech	• Provide resources in terms of professional development and supplies • Provide additional teacher support and materials management • Participate in relationships with the teachers and administration

Agent-based model rules: The change equations developed to model this case study follow the structure provided in Chapter 4. The values of internal and external weight used for School 1 are 0.6 and 0.4 respectively, and the same for School 2 are 0.4 and 0.6 respectively. These weights are estimated based upon the scale of the intervention and the external trends discussed before for the two schools. Another important parameter in the change equation is the transient phase parameter K which defines the length of time until a new steady-state is reached. For this case study, K is taken as 0.1, which means that the length of time for transition from transient phase to complete steady state phase was assumed to be about 46 months. However, as Georgia Tech provides professional development to the teachers every year, certain attributes like teacher's inquiry teaching re-enter transient state at the start of each year of the intervention. The different types of change equations used in the SLIDER case study, which are change in attributes, effect of resource flows, and change in relationships, are presented below.

Rules: Change in Attributes

As discussed during the framework development in *Section 4.3,* there are two components in the internal change equation term ($p_{internal}$): the transient and the steady state components. Change in the attributes of an agent is affected by the agent's own attributes, the relationships it has with the other agents, and the other agents' attributes. Following that methodology, the internal change equation components representing the change in the **support for intervention** of a teacher can be represented using the Equations (24) and (25) below.

$$p_{transient\_support}(t+1) = w_{teacher} \cdot \underbrace{teacher(t)}_{attributes} + w_{admin} \cdot \underbrace{relationship(t)}_{teacher:admin} \cdot \underbrace{admin(t)}_{support} \quad (24)$$

$$p_{steady\_support}(t+1) = w_{teacher} \cdot \underbrace{I_{teacher}(t)}_{attributes} + w_{admin} \cdot \underbrace{relationship(t)}_{teacher:admin} \cdot \underbrace{I_{admin}(t)}_{support} \quad (25)$$

Where $I_x(t) = \begin{cases} +1; \text{if } x(t) - x(t-1) > 0 \\ 0; \ \text{if } x(t) - x(t-1) = 0 \\ -1; \ \text{if } x(t) - x(t-1) < 0 \end{cases}$

In the Equations (24) and (25), $p_{transient\_support}(t+1)$ and $p_{steady\_support}(t+1)$ represent the transient and steady state change equation components for teacher's support for intervention in time period $t+1$ respectively. $I_x(t)$ is a modified indicator function which follows the logic as represented by the equation, and x is a given attribute. The first term, in both transient and steady state equations, is for the teacher's own attributes and $w_{teacher}$ is the weight for this term. The second term is for the relationship of the teacher with the administration, where w_{admin} is the weight for this term, $\underbrace{relationship(t)}_{teacher:admin}$ is the relationship between the teacher and administration at time t, and $\underbrace{admin(t)}_{support}$ is the support for intervention of the administration at time t. $w_{teacher} + w_{admin} = 1$; and the weights can be tuned for the particular case. In this case study, the values used for $w_{teacher}$ and w_{admin} are 0.1 and 0.9 respectively. In the above equations and the equations presented next, the function $p(.)$ can take both positive and negative values. If $p(.)$ is positive, then it represents the probability of the attribute to improve ($p_{improve}$), and $1 - p(.)$ represents the probability of the attribute to stay the same (p_{stay}). p_{worsen} in this case is zero. But, if $p(.)$ is negative, then the absolute value $|p(.)|$ represents the probability of the attribute to worsen (p_{worsen}), and $1 - |p(.)|$ represents the probability of the attribute to stay the same (p_{stay}). $p_{improve}$ in this case is zero. The above equations are representative of the change equations for various agent attributes modeled.

Rules: Effect of Resource Flows

The resources flow taking place from Georgia Tech to the school agents, specifically in the form of professional development, affected teachers' inquiry teaching skill, and content knowledge. A teacher's inquiry teaching skill is also affected by the relationships that teacher has with the other teachers in the school, especially if they teach the same subject and in the same grade level. In the example illustrated in Equations (26) and (27) below, there are two 8th grade physical science teachers who are receiving professional development from Georgia Tech, $teacher1$ and $teacher2$, and we are looking at the change in inquiry teaching skill of $teacher1$.

$$p_{transient\_inq}(t+1) \; = \; w_{teacher2} \cdot \underbrace{rel(t)}_{teacher1:2} . \underbrace{teacher2(t)}_{inquiry} \; + \; w_{pd}.GT_{pd}(t) \cdot \underbrace{teacher1(t)}_{self-eff\,\&\,sup} \tag{26}$$

$$p_{steady\_inq}(t+1) \; = \; w_{teacher2} \cdot \underbrace{rel(t)}_{teacher1:2} . \underbrace{I_{teacher2}(t)}_{inquiry} \; + \; w_{pd}.GT_{pd}(t) \cdot \underbrace{I_{teacher1}(t)}_{self-eff\,\&\,sup} \tag{27}$$

Where $I_x(t) = \begin{cases} +1; \text{ if } x(t) - x(t-1) > 0 \\ \;0; \;\text{ if } x(t) - x(t-1) = 0 \\ -1; \text{ if } x(t) - x(t-1) < 0 \end{cases}$

Here, $p_{transient\_inq}(t+1)$ and $p_{steady\_inq}(t+1)$ represent the transient and steady state change equation components for teacher's inquiry teaching in time period $t+1$ respectively. $I_x(t)$ is a modified indicator function which follows the logic as represented by the equation, and x is a given attribute. The first term, in both transient and steady state equations, is for the relationship between $teacher1$ and $teacher2$, $w_{teacher2}$ is the weight for this term, $\underbrace{rel(t)}_{teacher1:2}$ is the relationship between $teacher1$ and $teacher2$ at time period t, and $\underbrace{teacher2(t)}_{inquiry}$ is the inquiry teaching level of $teacher2$ at time period t. The second term is for the professional development being given to the teacher, w_{pd} is the weight for this term, $GT_{pd}(t)$ is the professional development level at time period t, and $\underbrace{teacher1(t)}_{self-eff\,\&\,sup}$ is the average of the self-efficacy and support for intervention of $teacher1$ at time period t. Again, $w_{teacher2} + w_{pd} = 1$. In this case study, the values used for $w_{teacher2}$ and w_{pd} are 0.1 and 0.9 respectively. The above equation follows the overall structure of the change equations where change in an agent's attributes is affected by three things: its own attributes, its relationship with the other agents, and the attributes of the other agents. The only difference is that, in the second term there is an additional factor multiplied to the agent's own attributes, which is $GT_{pd}(t)$. This is to capture the fact that there can be differences in the professional development being given during different interventions as well as during different times of the intervention, some might be better designed than others, some might have more resources to spend on professional development, and so on.

Rules: Change in Relationships

For this intervention, we do not model the social network effects in the students, like the EWB case study presented in Chapter 3, since we model the whole class as an agent. However, we have modeled the change in relationships amongst the teacher, administration, and Georgia Tech. Since these relationships are not peer relationships, the way in which these change have to be modeled are different than the social network concepts used earlier, namely *homophily* and *structural balance*. For example, the relationship between administration and Georgia Tech is affected by the support for intervention that the administration has during the intervention. For relationships, it is assumed that they start off at an initial level, and then any change in the attributes affecting the relationship would cause a change in that relationship. So there is no transient phase term in this case. Equation (28) represents the change in relationship between administration and Georgia Tech.

$$p_{change\_rel\_admin\_GT}(t+1) \;=\; w_{admin} \cdot \underbrace{I_{admin}(t)}_{support} \tag{28}$$

Where $I_x(t) = \begin{cases} +1; \text{if } x(t) - x(t-1) > 0 \\ \;0;\;\text{ if } x(t) - x(t-1) = 0 \\ -1;\;\text{if } x(t) - x(t-1) < 0 \end{cases}$

In Equation (28), $p_{change\_rel\_admin\_GT}(t+1)$ represents the internal change equation for the relationship between administration and Georgia Tech in time period $t+1$ respectively. $I_x(t)$ is a modified indicator function which follows the logic as represented by the equation, and x is a given attribute. This relationship is affected by the attribute support for intervention of the administration and w_{admin} is the weight for this term, which was taken as 1 for this case study.

Conceptual model validation: The conceptual model for the SLIDER intervention at the individual schools is complete now. Before moving forward, at this stage it is critical to validate the conceptual model using subject matter experts (SMEs). SMEs chosen to validate the conceptual model were the design and implementation team of SLIDER, public policy researchers, as well as one of the teachers involved in the intervention. There were a total of about 10 SMEs used to verify the conceptual model. Note that this process was actually done even concurrently while the conceptual model was being built. The conceptual model thus went through a number of iterations based upon the SMEs feedback and the conceptual model presented above is the final version incorporating all of this feedback.

***Data Collection*:** Data was collected for the schools involved in the intervention. This was done to both populate the initial state of the simulation model and then to later validate the results of the simulation model. Data was already being collected by the Georgia Tech SLIDER team throughout the implementation of the intervention. This data included details about the demographics in the schools, the experience and content knowledge of teachers, classroom observations, student pre-post SLIDER test scores, the number of students in each class, the support for intervention of the teachers and administration etc. Data was also collected about the trends at the schools through the data made publicly available by the Georgia Department of Education and School Digger [71, 72]. Any remaining data that was needed, either for the initial state or to validate the model, was acquired through surveys given to the SLIDER team at Georgia Tech, and the teachers and administration at the schools involved in the intervention. Apart from populating the initial states of the simulation model (at time t = 0, at the start of the intervention), the data collected from the schools also helps in estimating the weights and other parameters used in the change equations (via the scale of the intervention and external trends). Once these parameters are estimated, the state transition probabilities are calculated using the change equations, and the change in the system states is mapped over the duration of the intervention. This change in the system, along with the final state of the attributes and relationships, is compared with the observed change in the system (through the data collected) to validate the usability of the model.

Computer simulation model: The agent-based simulation model was built using the object-oriented programming language *C#* in Microsoft Visual Studio 2010. The system configuration was Windows 2007, 64 bits, 8GB Ram and 1.73 GHz Processor. An object-oriented programming language was chosen because it allows the agent classes to be modeled as different classes in the program and agents as objects in the program. It also provides flexibility to implement the various agent-based model rules developed for this intervention and other modeling features. For example, the flexibility of having changing time steps during the time horizon of the intervention could be easily implemented in *C#*. The computer model built is verified using various techniques described in the model validation phase of the framework. This completes the model design phase.

5.2.3 Model Analysis

Simulation results: Since this is a stochastic model, first the number of samples have to be determined such that the internal variance in the model is negligible compared to the mean of the outputs. Table 13 shows the variance in the overall fitness output for different number of runs. Fitness can vary from $(0, \infty)$.

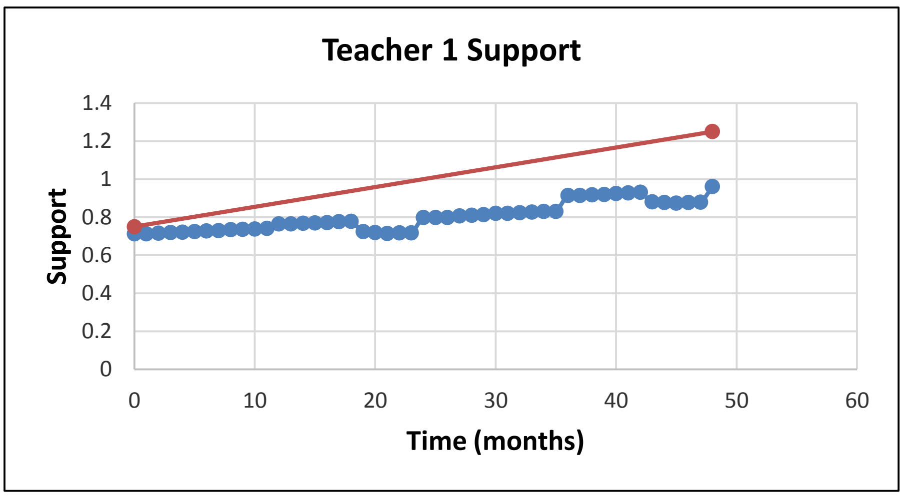
Table 13: Mean and variance in the final fitness output

Number of samples	Number of runs	Final Fitness		Run time (min:sec)
		Mean	Std_dev	
100	100	1.33	0.054	00:20.89
1000	100	1.33	0.051	03:20.73
100	1000	1.33	0.018	03:21.65
1000	1000	1.33	0.018	32:02.43

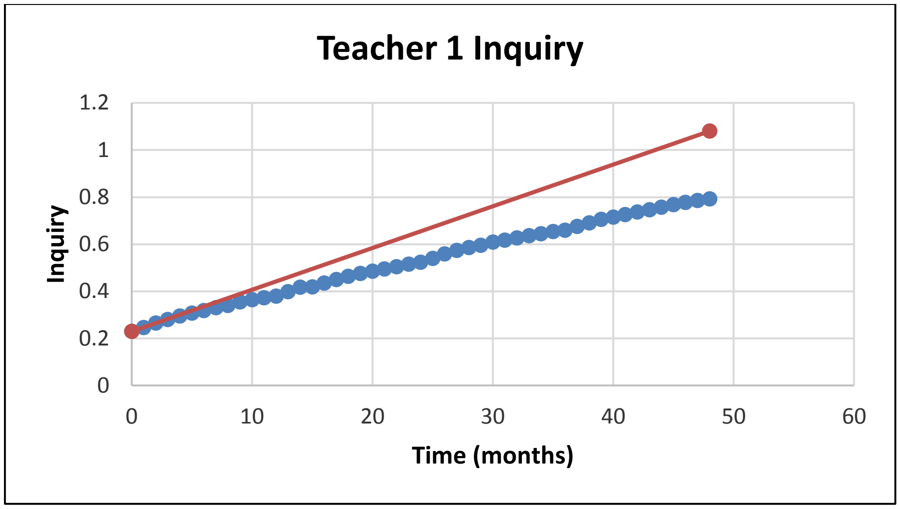
In the above table number of runs is the parameter being decided and number of samples if the number of times each sample of the number of runs is taken in order to calculate the standard deviation. From the above table it can be seen that the standard deviation for 1000 runs is two orders of magnitude less than the mean. Hence, each model run will consist of 1000 simulation runs, such that the model output is almost deterministic. Decreasing the internal variance in the model outputs is critical for conducting sensitivity analysis and Method of Morris experiment during the model analysis phase. Next, the simulation results are presented for the two schools for some of the key attributes considered in the intervention.

School 1

The change in teacher, classroom, and administration attributes is presented next. There were 2 teachers in School 1 which were involved in the SLIDER intervention. For brevity only Teacher 1's results are shown below in Figure 16. Teacher 2's results are presented in Appendix A. The change in the three gaps and the fitness values are also shown below. The results are presented for the first four years of the intervention. In all the graphs shown below, the blue line represents the model results over time, and the red line represents the actual measured values for these attributes. The red curve ties to the model validation phase of the framework, where the simulation results are validated using data.

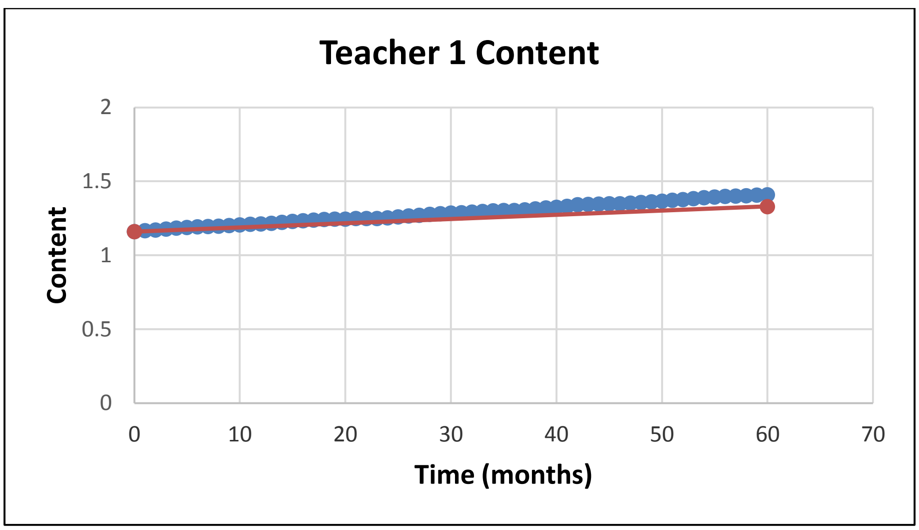


(a)



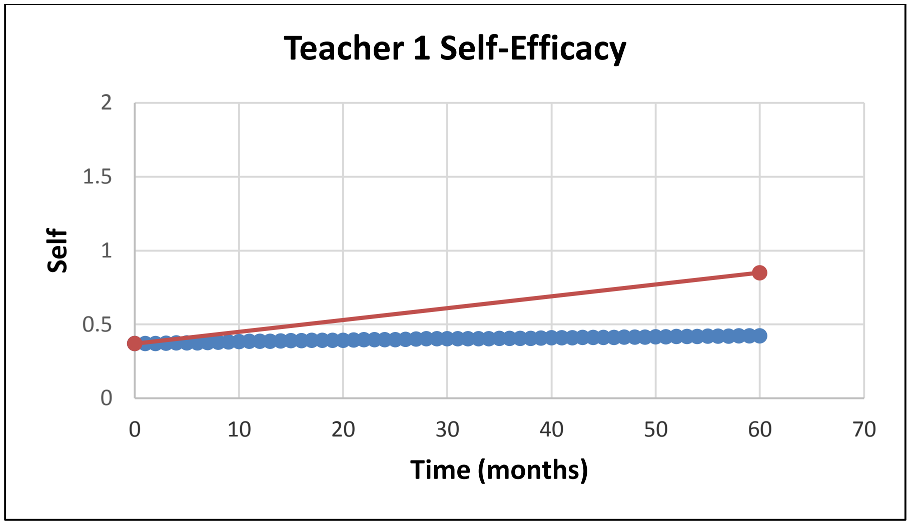
Teacher 1 Inquiry
Inquiry
1.2
1
0.8
0.6
0.4
0.2
0
0
10
20
30
40
50
60
Time (months)

(b)



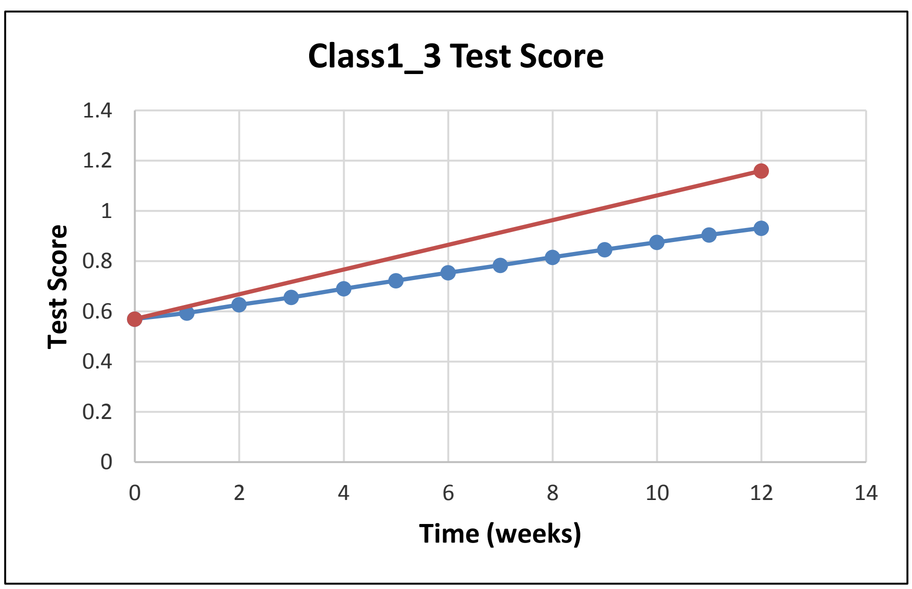
Teacher 1 Content
Content
2
1.5
1
0.5
0
0
10
20
30
40
50
60
70
Time (months)

(c)

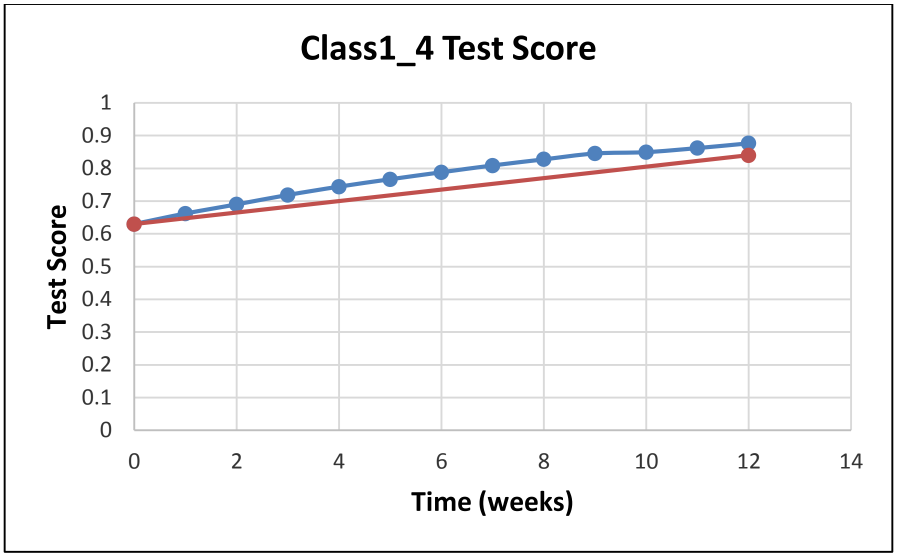


(d)

Figure 16: (a) Change in Teacher 1's support for intervention; (b) Change in Teacher 1's inquiry teaching skill; (c) Change in Teacher 1's content knowledge; (d) Change in Teacher 1's self-efficacy



(a)



(b)

Figure 17: (a) Change in Class 1's SLIDER test scores in year 3; (b) Change in Class 1's SLIDER test scores in year 4

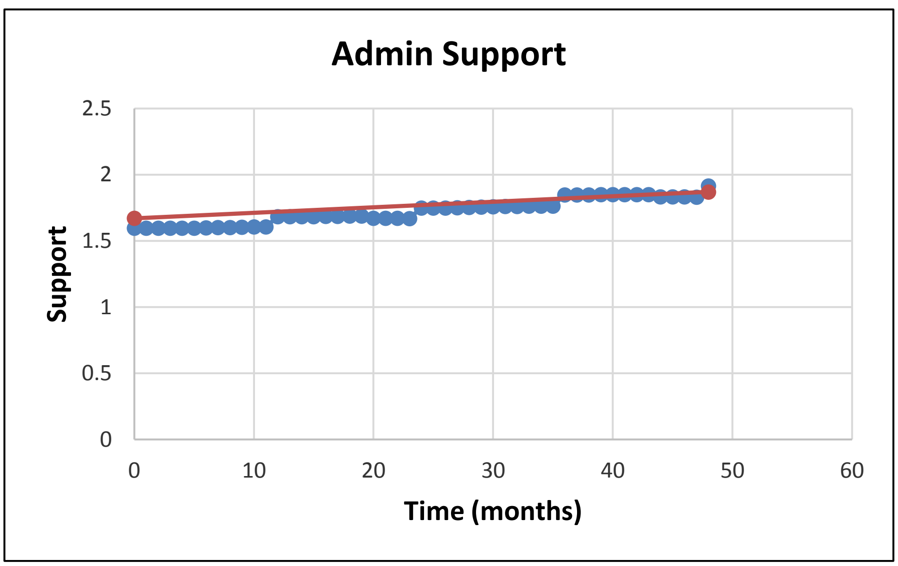
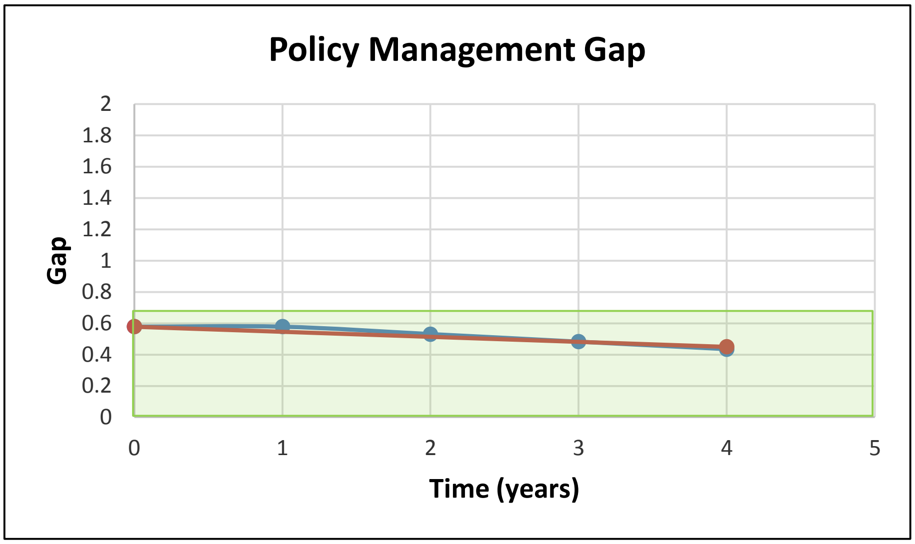
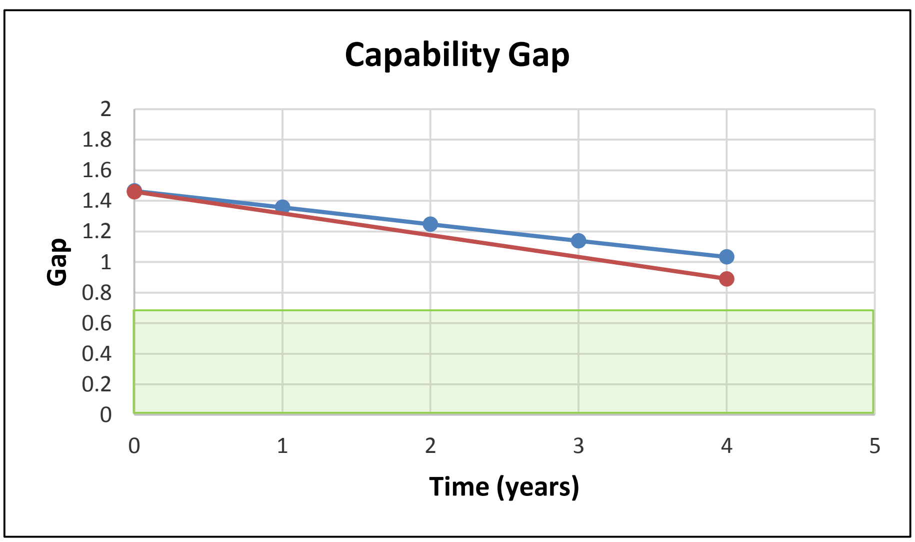


Figure 18: Change in Administration's support for intervention

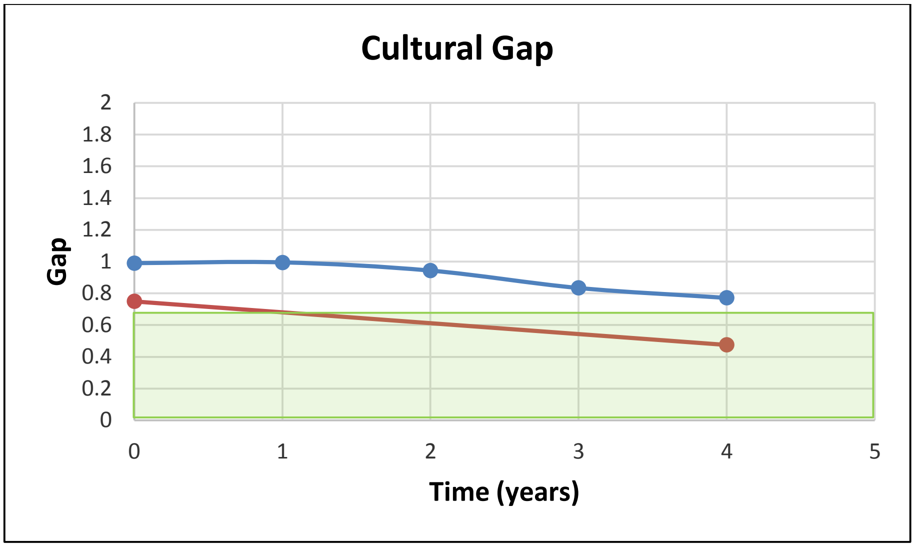
It can be seen from the above change in attributes, that the system state of School 1 is improving and this is consistent with the data collected for this particular school. The three gaps and fitness are presented next.



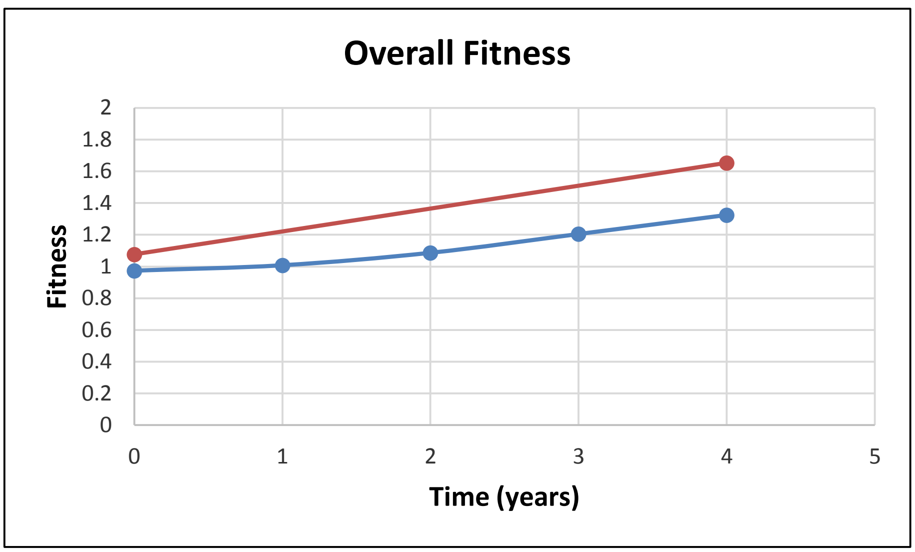
(a)



(b)



(c)



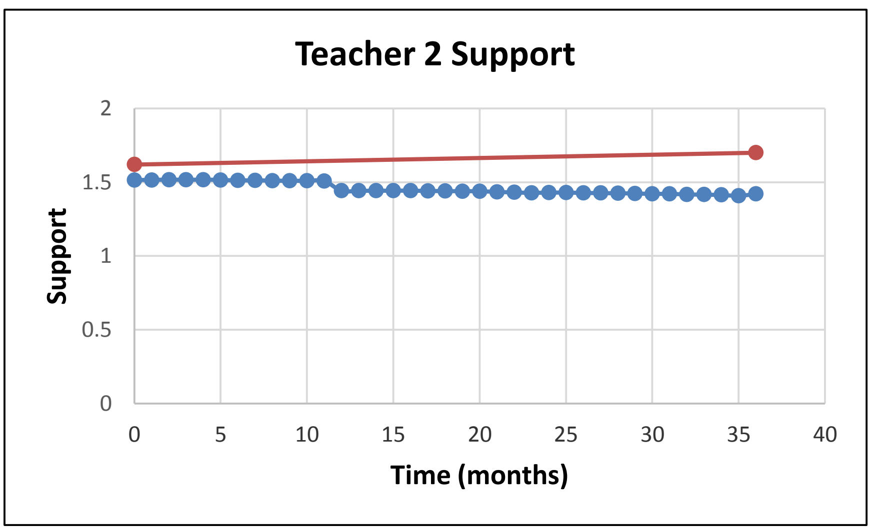
(d)

Figure 19: (a) Policy management gap in School 1; (b) Capability gap in School 1; (c) Cultural gap in School 1; (d) Overall fitness of School 1

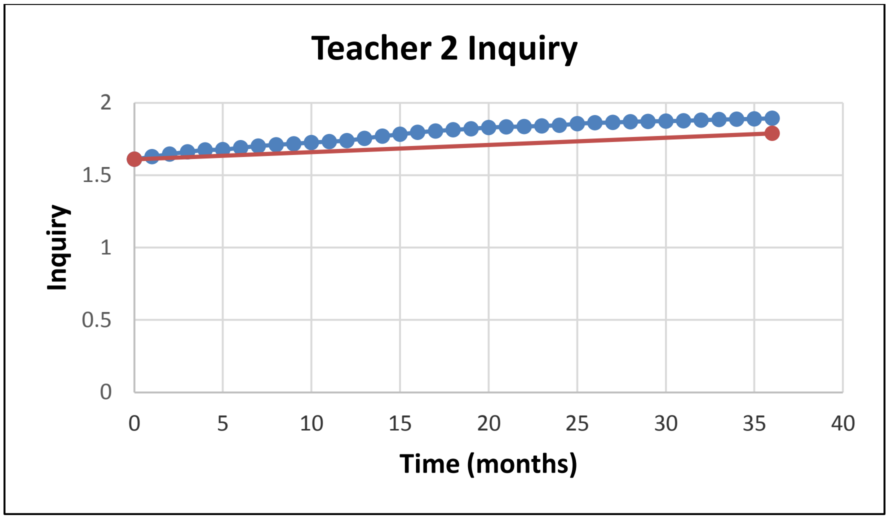
It can be seen from the above graphs that the three gaps at School 1 are decreasing and hence the overall fitness is improving. The shaded region with green boundary represents the region of acceptable zone. For all three gaps, School 1 is either close to or within the acceptable zone by the end of the fourth year (blue line). This is consistent with the actual data as well (red line).

School 2

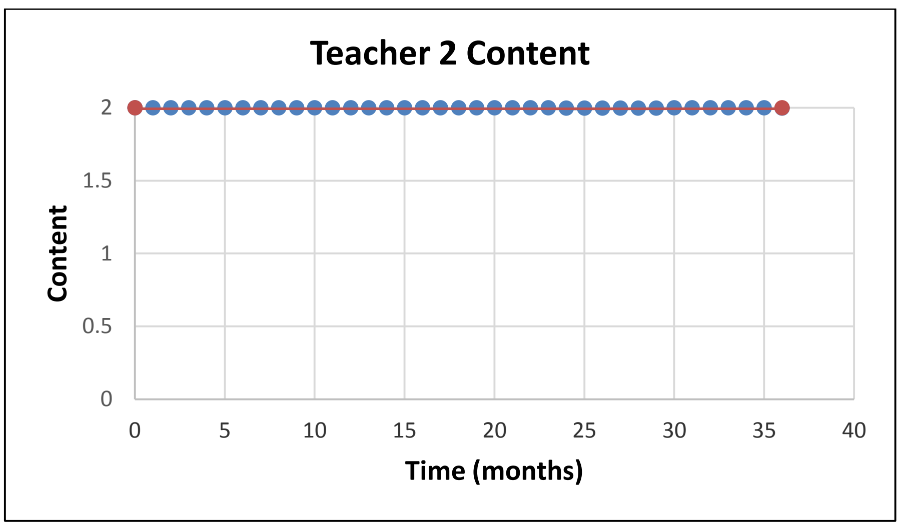
Similar to School 1, the changes in teacher, classroom, and administration attributes are presented next for School 2. There were 3 teachers in School 2 which were involved in the SLIDER intervention. Only Teacher 2's results are presented below, Teacher 1's and Teacher 3's results are presented in Appendix A. Change in the three gaps and the fitness values are shown below as well, in Figures 20-23. The results are presented for the first four years of the intervention. In all the graphs shown below, the blue line represents the model results over time, and the red line represents the actual values for these attributes.



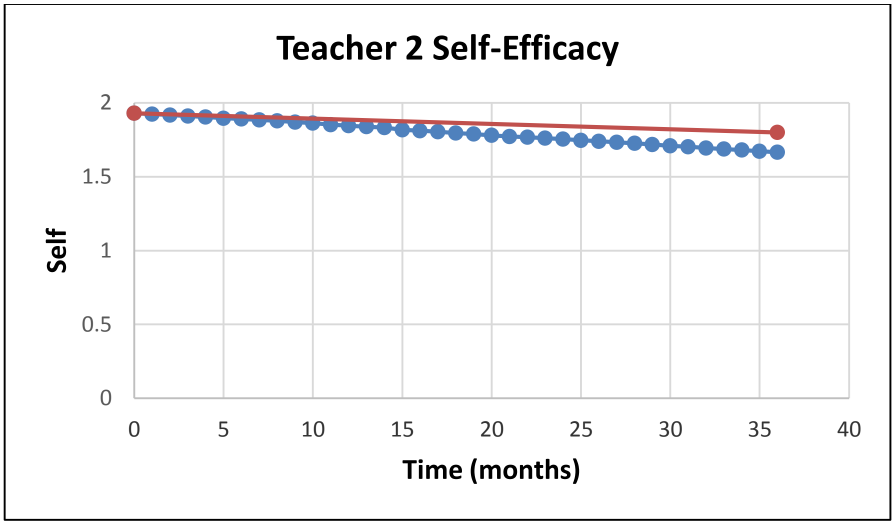
(a)


Teacher 2 Inquiry
Inquiry
2
1.5
1
0.5
0
0
5
10
15
20
25
30
35
40
Time (months)

(b)

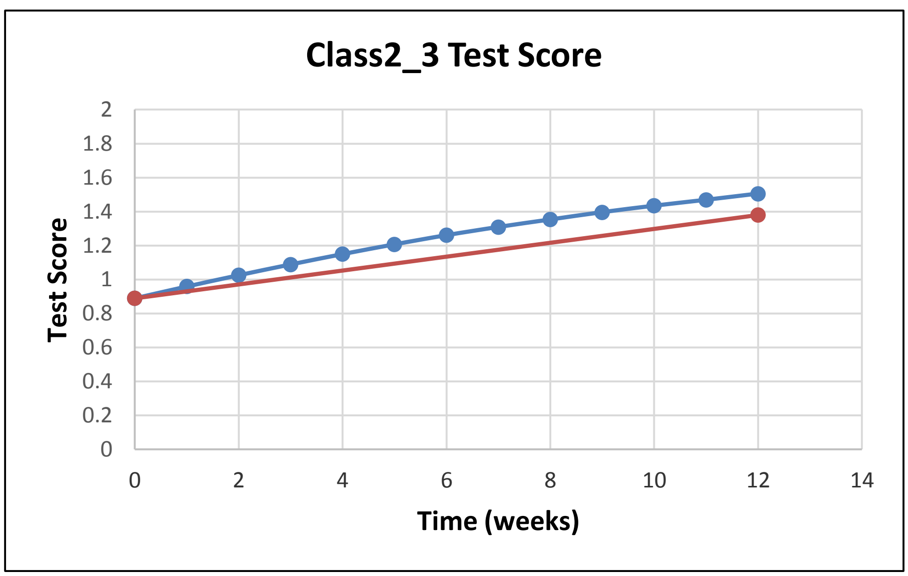

Teacher 2 Content
Content
2
1.5
1
0.5
0
0
5
10
15
20
25
30
35
40
Time (months)

(c)

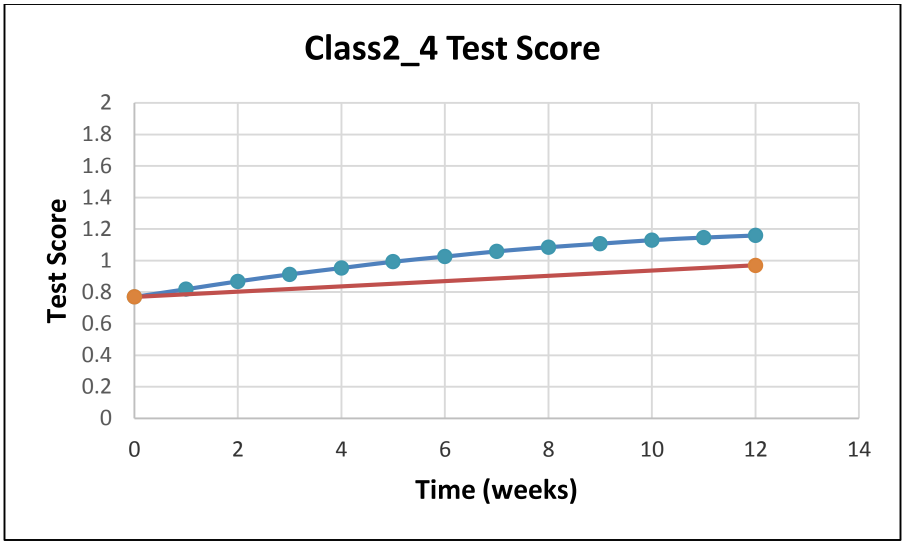


(d)

Figure 20: (a) Change in Teacher 2's support for intervention; (b) Change in Teacher 2's inquiry teaching skill; (c) Change in Teacher 2's content knowledge; (d) Change in Teacher 2's self-efficacy



(a)



(b)

Figure 21: (a) Change in Class 2's SLIDER test scores in year 3; (b) Change in Class 2's SLIDER test scores in year 4

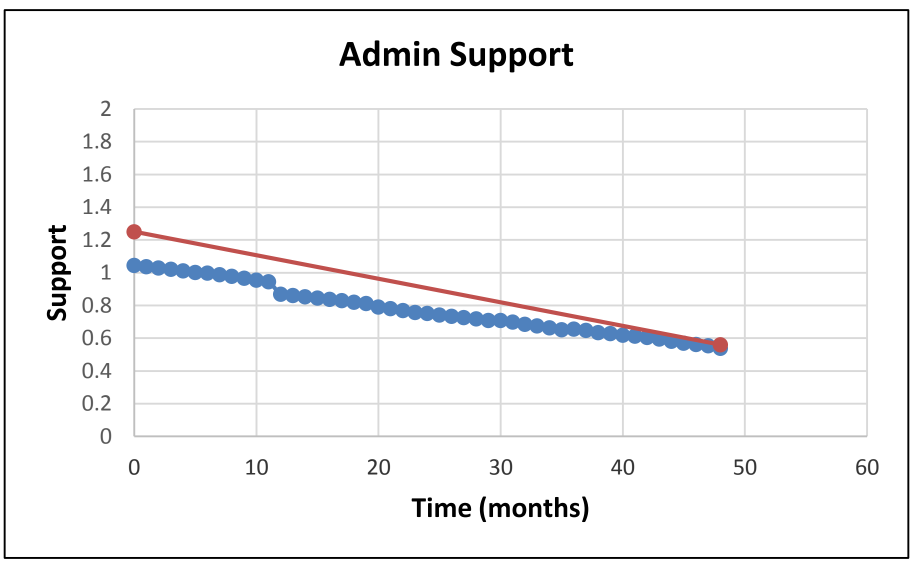
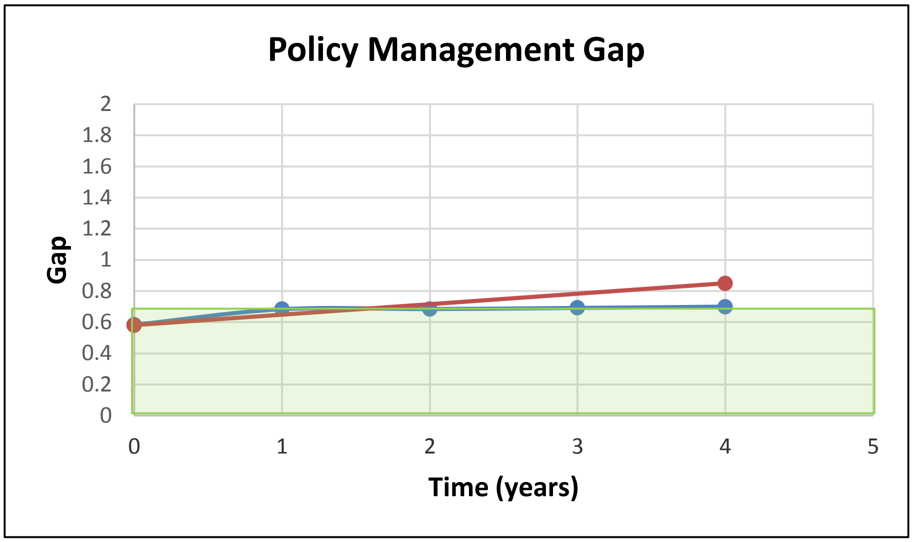
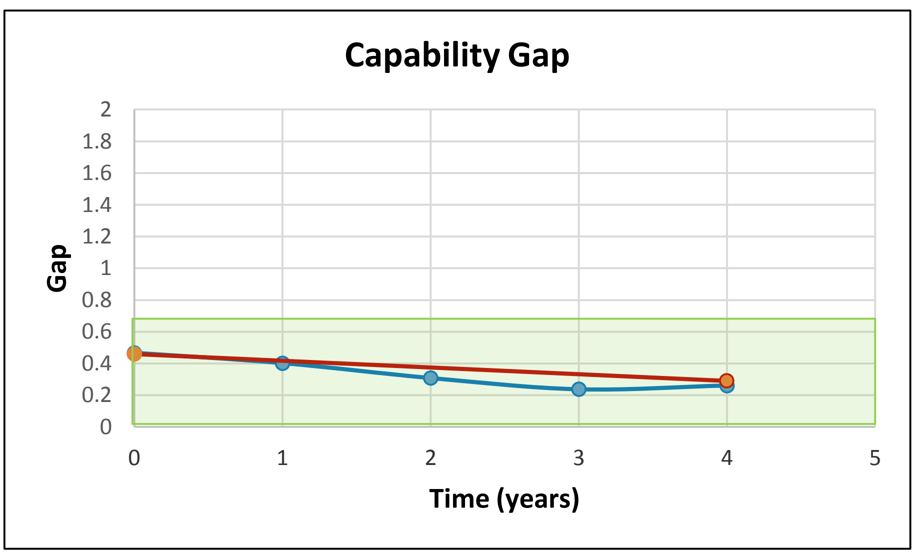


Figure 22: Change in Administration's support for intervention

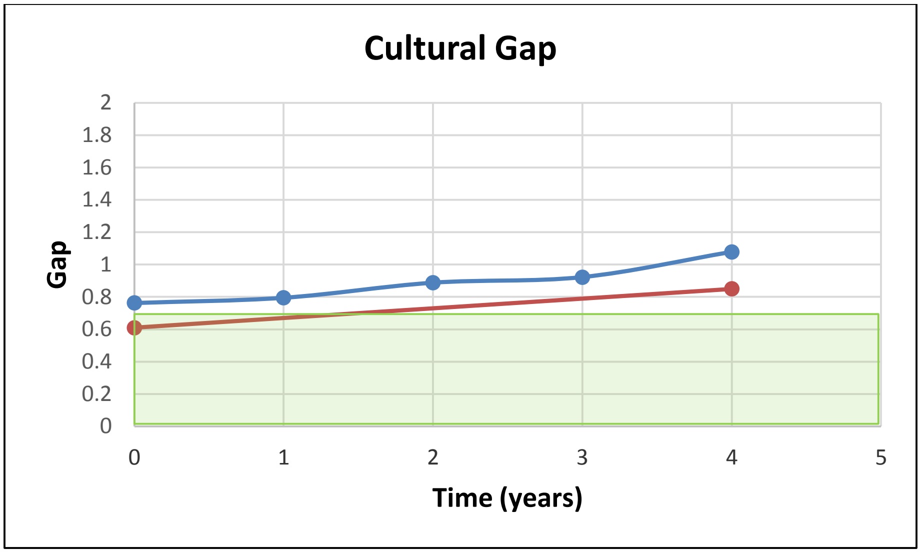
For School 2, it can be seen that the support for intervention of the teachers and the administration is going down. However, the teacher's inquiry teaching skills are improving. The three gaps and fitness are presented next to better analyze the change in the system state of School 2; whether it is moving closer to the acceptable zone or away from the acceptable zone.



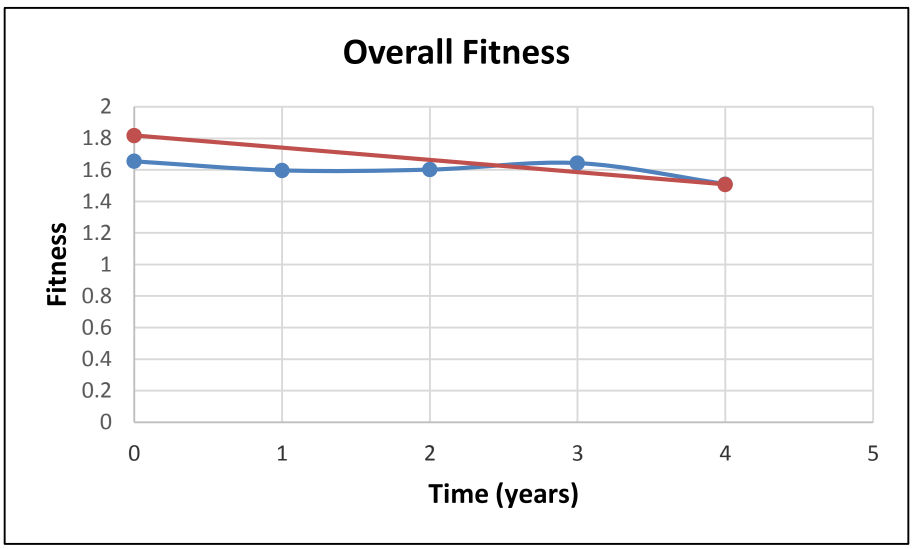
(a)



(b)



(c)



(d)

Figure 23: (a) Policy management gap in School 2; (b) Capability gap in School 2; (c) Cultural gap in School 2; (d) Overall fitness of School 2

Apart from the capability gap, the remaining two gaps are worsening for School 2, and the overall fitness is also going down over the duration of the intervention. School 2 actually started off in a better state in comparison to School 1, however because of external trends and the agent attributes and relationships at School 2, the gaps didn't reduce and the fitness went down.

Table 14 also compares the probability of sustainability of the intervention, $P(I)$ for School 1 and School 2 across different years of the intervention. The method to calculate $P(I)$ was presented in *Section 4.2* of Chapter 4.

Table 14: Probability of sustainability of the intervention at School 1 and 2

Year	School	PM_gap	Cu_gap	Ca_gap	$P(I)$
0	School 1	0.58	0.99	1.47	0.27
1	School 1	0.58	0.99	1.36	0.33
2	School 1	0.53	0.94	1.25	0.40
3	School 1	0.48	0.83	1.14	0.48
4	School 1	0.43	0.77	1.03	0.55
0	School 2	0.58	0.76	0.47	0.72
1	School 2	0.68	0.80	0.40	0.67
2	School 2	0.68	0.89	0.31	0.60
3	School 2	0.69	0.94	0.24	0.58
4	School 2	0.70	1.11	0.26	0.47

It can be seen from the above table that the probability of success of the intervention is increasing for School 1 whereas it is decreasing for School 2. Recall that the derivative of the change in the gaps also goes as an input into the calculation of $P(I)$. $P(I)$ is an intuitive measure to understand the chances of sustainability and risks involved in intervening in the particular system.

Sensitivity Analysis: An important step in the model analysis phase is sensitivity analysis. This is conducted upon the parameters of the change equations. Later, the sensitivity analysis of inputs on the outputs via the Method of Morris will be presented. The main purpose of doing sensitivity analysis on the parameters is to check if the model results are highly sensitive to a particular parameter (in which case extreme care must be paid when choosing the values of those parameters). Table 15 shows how the fitness function value changes as the internal and external weights change in the change equations.

Table 15: Sensitivity analysis with respect to internal and external weights

w_int	w_ext	Fitness
0%	100%	1.072
10%	90%	1.143
20%	80%	1.247
30%	70%	1.366
35%	65%	1.427
36%	64%	1.426
37%	63%	1.459
38%	62%	1.446
39%	61%	1.488
40%	60%	1.486
41%	59%	1.497
42%	58%	1.521
43%	57%	1.540

Table 15 continued

44%	56%	1.553
45%	55%	1.576
50%	50%	1.630
60%	40%	1.740
70%	30%	1.857
80%	20%	1.949
90%	10%	1.967
100%	0%	2.011

From the above table it can be seen that change in the internal and external weights does not change the fitness values drastically. The above fitness value is for School 2, and the values of internal and external weight used for this school were 0.4 and 0.6 respectively. Note that as the internal weight is increased, and consequently the external weight is decreased, the fitness value goes up as more weight is assigned to the effects taking place through the intervention.

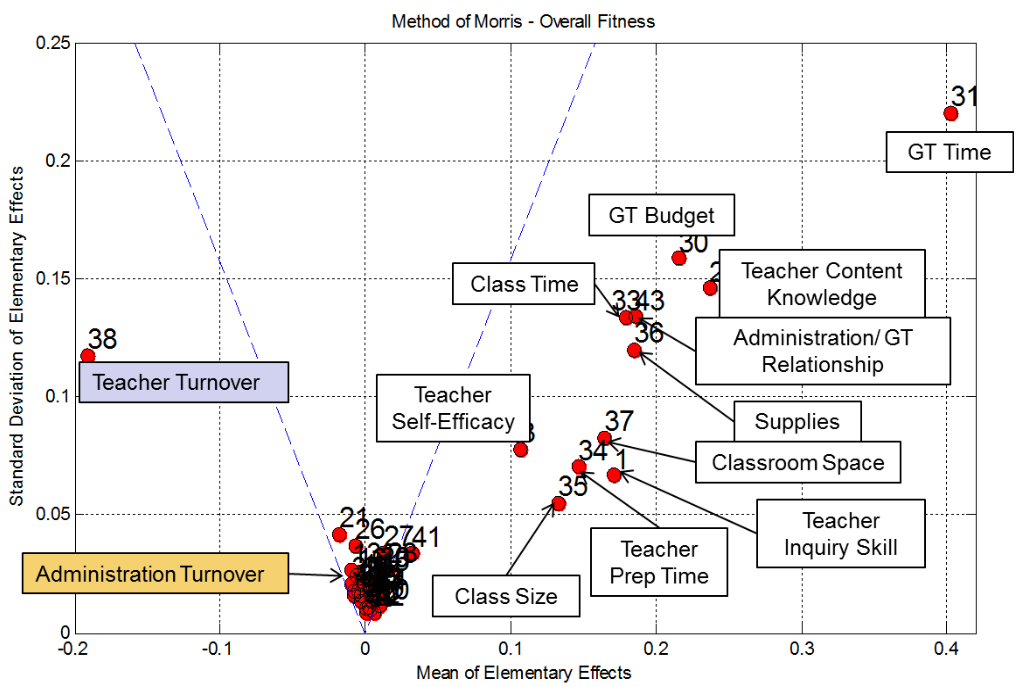
Another important parameter in the change equation is the transient phase parameter K. K defines the length of time until a new steady state is reached. For the SLIDER case study, K was taken as 0.1, which means that the length of time for transition from transient phase to complete steady state phase was assumed to be about 46 months. However, as Georgia Tech provides professional development to the teachers every year, certain attributes like teacher's inquiry teaching re-enter transient state at the start of each year of the intervention. Table 16 presents the change in the overall fitness as K changes.

Table 16: Sensitivity analysis with respect to the transient phase parameter

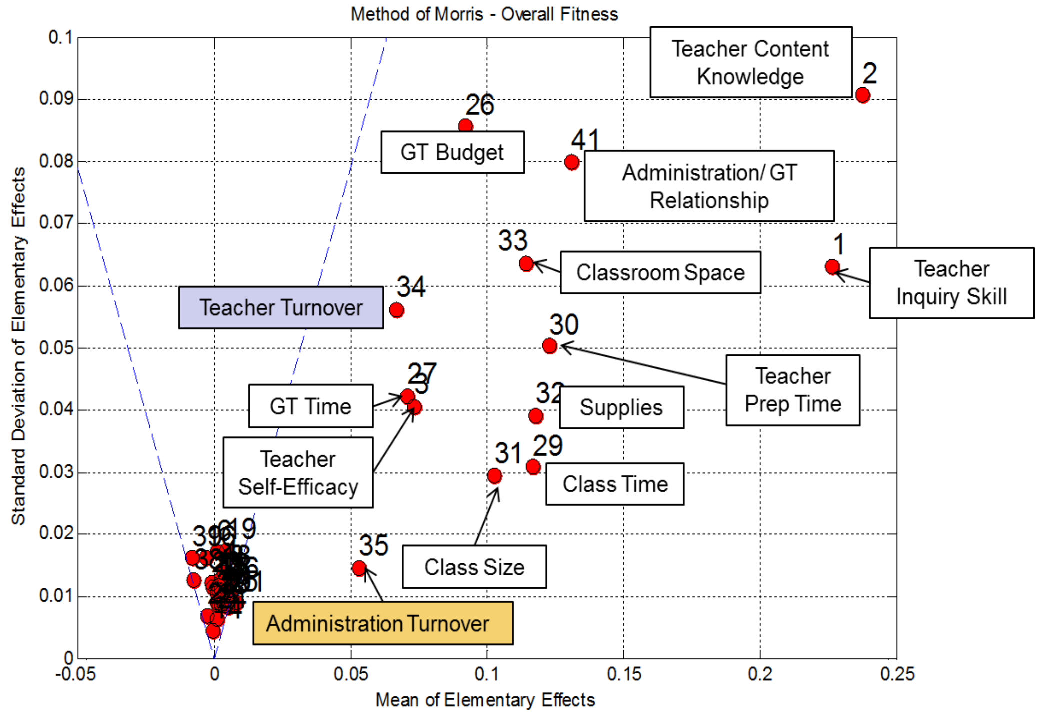
Time until Steady State (months)	***K***	**Fitness**
92	0.05	1.625
77	0.06	1.563
66	0.07	1.534
58	0.08	1.508
51	0.09	1.486
46	0.1	1.482
42	0.11	1.445
38	0.12	1.444
35	0.13	1.435
33	0.14	1.387
31	0.15	1.356

Similar to the previous result regarding the weights, from the above table it can be seen that the fitness value is not highly sensitive to the K value or the time until steady state. As the time until steady state increases, which means that the transient phase due to the intervention is longer, the fitness value increases as the intervention is able to have a higher impact. Hence, from this analysis it can be seen that the change equation parameter values don't change the model results drastically.

Method of Morris: The Method of Morris experiment is conducted for the two schools with 43 different input variables, including attributes of the teachers, administration, students, school environment, and Georgia Tech and the relationship modeled between the different actors at the start of the intervention. The response variable used to analyze the effect of the inputs is the overall fitness at the end of the intervention. The figures below shows a graph of the mean and standard deviation of the effect of inputs on the overall fitness of School 1 and School 2. The mean of the effects is the x-axis and the standard deviation of the effects is the y-axis. The dotted lines correspond to the equation: Mean = ±2·SEM, where SEM is the standard error of the mean and is equal to the standard deviation divided by the square root of number of random orientations for each input, which for this analysis is taken as 10. The circles represent the different inputs. Inputs laying outside of the ‘v’ or far from zero have an impact on the response variable that is statistically different than zero.



(a)



(b)

Figure 24: (a) Barriers & enablers for School 1; (b) Barriers & enablers for School 2

From the above figures, it can be observed that the major factors that can lead to a successful implementation of this intervention are the teacher's attributes with respect to inquiry teaching, content knowledge, and self-efficacy, the relationship between the administration and Georgia Tech, school environment attributes likes class duration, classroom space, availability of supplies, and teacher preparation time, and time spent by Georgia Tech on professional development. These are the common attributes obtained for both schools, and which have a positive impact on the overall fitness. However, the attribute teacher turnover rate has a negative impact on the overall fitness in School 1. This makes sense as teachers are the ones receiving professional development and implementing the intervention and if there is a high turnover rate then sustaining this intervention will not be feasible. However, in comparison at School 2, the teacher turnover rate and the administration turnover rate have a positive impact on the overall fitness of School 2. This is a counter-intuitive insight and it can be explained via the phenomenon that the attributes of the agents in School 2 are going down over time, and when there is a turnover of these agents, the attribute levels again increase. The assumption here is that when the turnover happens, the agents which replace the current agents have attribute levels close to the current agents at the time of the start of the intervention. However, this may not always be the case as it might be difficult to attract good talent to the school if the school performance is going down.

MoM experiment was also conducted using the probability of sustainability function, just to compare its results with the MoM experiment using the overall fitness function. As can be seen from the MoM graph presented below for School 1, the set of important attributes obtained at the end were the same for these two cases.

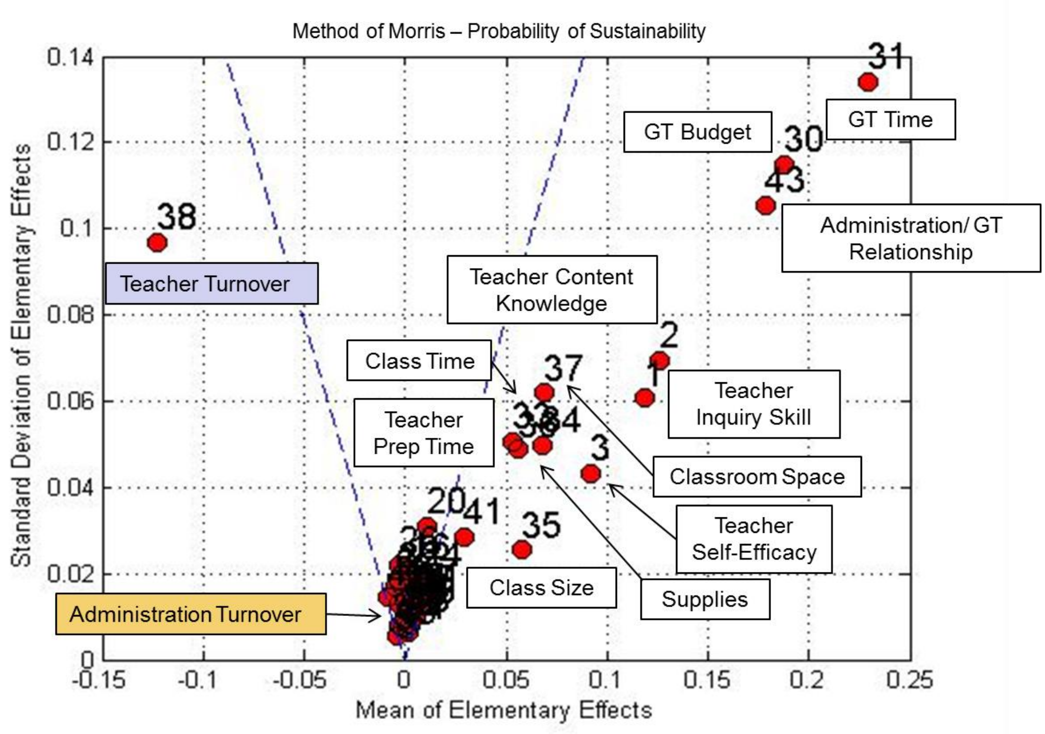


Figure 25: MoM – Probability of sustainability function

This shows that the MoM result are quite robust with respect to the different functional forms of the fitness function. Hence, the overall fitness function can be used over the probability of sustainability function, to conduct the MoM experiment, as it is a simpler function. This completes the model analysis phase of the framework.

5.2.4 Model Validation

The following combinations of steps, from among the steps discussed in the model validation phase of the framework development in *Section 4.5*, were completed to validate the model built for this case study. These steps have been applied across various phases of application of the framework to build the model for this case study.

Conceptual model and face validation: In order to complete this step, 10 subject matter experts (SMEs) were enlisted to test whether the model and its behavior were conceptually logical and whether the model's input-output relationships were reasonable. These subject matter experts were school teachers, educational researchers, and public policy researchers. They tested the model for completeness, consistency, coherence, and correctness as described in the *4C's* framework during the model validation phase of the framework in Chapter 4. Since SMEs were used at various steps of the model building process, this helps in developing a model that is easier to validate at the end.

Animation/graphics: As shown in the simulation results section of the model analysis phase, the model's operational behavior is displayed graphically as the model variables move through time. This behavior is shown to the SMEs and tested for validation. Also, the change in the model variables is compared to the data collected as described next in the data validation step.

Data validation: The changes in the attributes of the agents at School 1 and School 2 are compared to the data collected with respect to these attributes. Overall, model results were coherent with the changes taking place in the attributes of the school agents. In some instances, this exercise led to insights about changes that could be made in the parameters of the change equation of that attribute. For example, for the change in class test scores in School 2, the model change was coming higher than the actual values observed. This led to the incorporation of a slight negative trend (modeled by the external part of the change equation) in the class test scores.

<u>Degenerate and Extreme Condition Tests</u>: The model output is analyzed for extreme values of the input parameters. From the sensitivity analysis section, it can be seen that for extreme values of the weights, w_{int} and w_{ext} in the change equations, the model behavior is plausible, i.e. as w_{int} becomes 100%, the overall fitness is the maximum and as w_{ext} becomes 100%, the overall fitness is the minimum. Also, for extreme values of the transient phase parameter, K, the overall fitness is maximum when K is minimum, this is because the transient effects of the intervention are at the maximum when K is at its minimum.

<u>Event validation</u>: One of the major events that took place during the intervention was Georgia Tech pulling out of School 2 at the end of year 4. This can be very clearly explained through the change in the gaps and fitness of School 2. The gaps were growing (worsening) over the course of the intervention and consequently the fitness was decreasing (also worsening). The school was moving away from the acceptable zone and hence the model in this case indicates that this intervention would not have been sustainable. This supports the decision taken by Georgia Tech to stop implementing the intervention at this school.

<u>Comparison to Other Models</u>: The model results and insights about the barriers and enablers of this intervention obtained are consistent with the educational research studies discussed in the literature review section about interventions in the education system. The factors mentioned there are also found to be important in this case study.

<u>Internal validation</u>: Several evaluations of the model were used to determine the internal stochastic variability in the model. The standard deviation in the results for 1000 model iterations was three orders of magnitudes less than the mean which implies model consistency.

Parameter variability – Sensitivity analysis: The model was run under different sets of parameter and input conditions and model outputs were analyzed. This is discussed in detail in the model analysis section.

5.3 Modeling at a Higher Scale

The ESIM framework has been applied to model the change in each school during the SLIDER intervention. This micro-behavior within the school can also be utilized to model the behavior of the system at a more macro-scale, where the schools are modeled together. Earlier the system boundary was within the school, now it is been extended to all the schools in which SLIDER is being implemented. At a macro-level, the agent network looks like this:

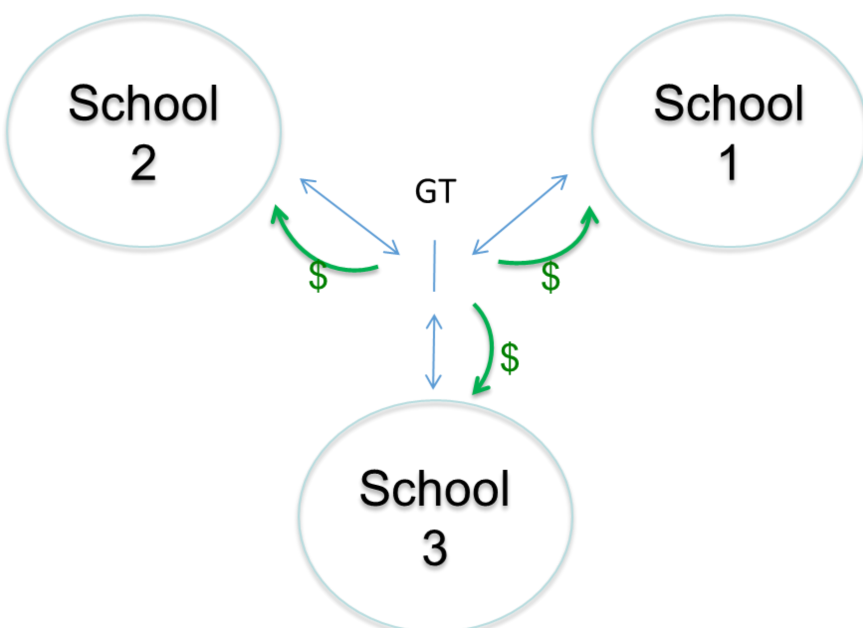


Figure 26: SLIDER network at a macro-level

Georgia Tech (GT), which is the intervening agency in the three schools, interacts with each school and invests resources in the form of time and money. Apart from the pre-determined professional development and materials supply, GT also provides additional teacher support and materials management support to the schools as and when it is required. Since School 3 is the control school with all three gaps close to zero, almost no additional support is required for this school. Teacher support is broken down into the following categories: emails, text/phone calls, video calls (Skype, Facetime), in person support, preparation for meetings, and emails. Materials management support is the support provided for managing and repairing the materials used in the intervention such as LEGO kits. For each of these categories, the total time spent by GT in School 1 and School 2 was obtained from the team. This is used to calculate the proportion of time spent by GT on additional support across the three schools.

The goal of modeling the schools together (at a macro-level) is to understand if the micro-behavior within the schools can explain the macro-behavior observed in terms of varying proportion of resource allocation from GT towards the three schools. From the micro-behavior at each school, gaps are known for each year at the three schools. Additional teacher support being provided is affected by the cultural and capability gaps, and additional materials management support being provided is affected by the policy management and capability gaps. It is assumed in the model that additional support being provided is proportional to the gaps which exist at the schools. Therefore, as School 3 has all three gaps as zero, no additional support is needed for School 3. For School 1 and School 2, the figure below shows the proportion of time spent by GT according to the model using the three gaps. This is compared with the actual proportion of time spent by

GT at School 1 and School 2. Since capability gap affects both teacher and materials management support, it is given a weight of 0.5 while calculating the proportion of time spent by GT in to the model. The remaining two gaps are given a weight of 0.25. Since no additional support was needed in Year 1 of the intervention, the results below are from Year 2 to 4.

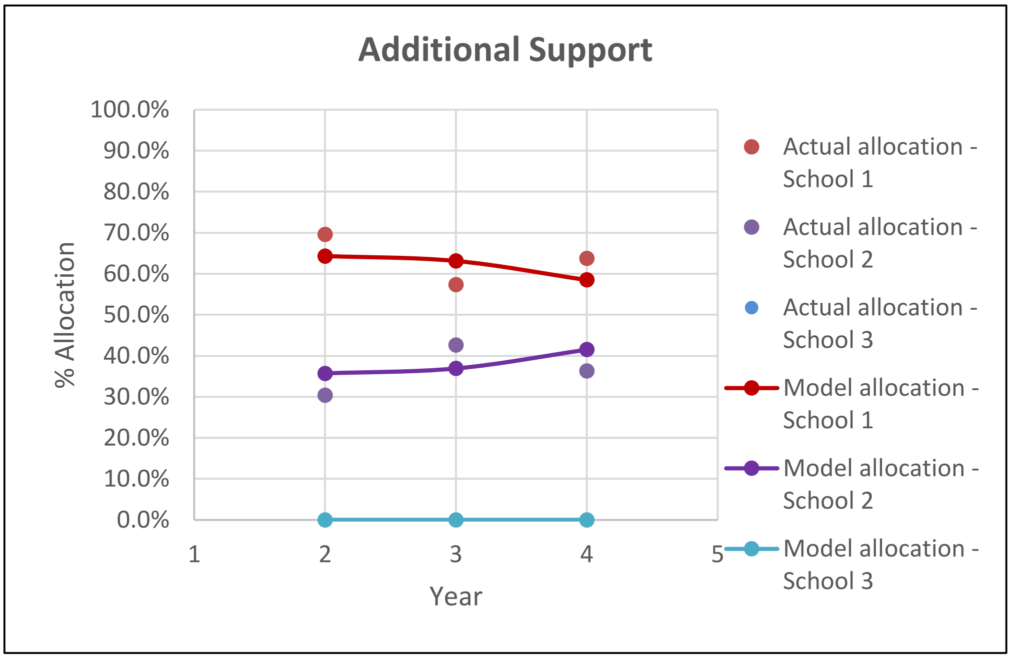


Figure 27: Proportion of additional support towards the three schools

From the above figure, it can be seen that the proportion of additional support that should be given to each of the three schools based upon the model results, and the proportion of actual additional support provided are quite close. The micro-behavior within each school, thus helps in explaining the macro-behavior, proportion of variable resource allocation from Georgia Tech towards the schools, when the schools are modeled together. This property of agent-based models is known as the *emergent property*, where the micro-behavior of the agents leads to a macro-behavior of the system. The above analysis demonstrates that the gaps have the potential to provide insights about the variable resource allocation from the school partner (Georgia Tech) towards the schools it is intervening in. However, this analysis should be conducted for more such case studies, where there are multiple schools across which the intervening agency has to allocate (limited) resources, to develop confidence in the use of gaps for making resource allocation decisions. In some cases, there might be differences between the intervening agency's resource allocation and what the gaps suggest. This would not necessarily mean that the gaps cannot guide the resource allocation process, it might also mean that the resource allocating agency did not allocate the resources optimally across the schools (i.e. based upon the needs of each school). Once confidence in the gap measures is developed for resource allocation purposes, this would also be a useful analysis to conduct when this framework is used to model interventions proactively, especially in a similar case where a single agency is allocating resources across multiple schools. This completes the application of the modeling framework to the SLIDER intervention. In the next chapter, major conclusions of this thesis and future research avenues are discussed.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

In this chapter, the main conclusions of this thesis, insights gained, and future research directions are discussed. First, discussion about the EWB case study is provided. Then, an overview of the ESIM framework is given. After that, discussion about the SLIDER case study is provided. Then, some key insights and policy implications of this work are presented, and finally, future avenues of research are discussed.

6.1 EWB Case Study

The *Engineers Without Borders* (EWB) case study was used in the development of the framework. An actual case study was chosen to develop the framework so that the framework developed and the techniques used in it are grounded in real world application. This case study was ideal for the development of the framework and as its first application because it had two different outcomes over two different years, which helped in developing insights about the success of this intervention. Also, the scale of this intervention was small enough to test the development and application of the framework. With the help of this case study, a more generalized framework was developed which is applicable across a broad range of education system interventions in the K-12 school system.

Two years of this intervention were modeled, one in which the intervention was successful and the other in which it was not. Collaboration with the lead teacher involved in this case study, and other educational researchers resulted in a model which was consistent with the intervention being modeled. The attributes that were identified to be most critical were the support for intervention of the principal and teacher, while the most important relationships were found to be between the principal and community, the teacher and community, and the teacher and school partner. Sensitivity analysis showed that the parameters used in the model are robust i.e. small changes in these parameters do not change the model outputs drastically. The Method of Morris experiment helped in quantifying the impact of the agents' attributes and relationships on the total money generated by the end of the intervention each year.

The modeling process for this case study confirmed the initial hypothesis that techniques from industrial engineering, systems engineering, and operations research can be used to model interventions in the education system. Agent-based modeling was used to simulate this case by representing the different actors involved in the intervention as agents and modeling the change in the agents' attributes via the relationships amongst agents and the resource flows taking place. Markov property was assumed while modeling change in the system states. Social network analysis was used to model the change in the relationships amongst the agents. Sensitivity analysis and Method of Morris was used to analyze the model results. Finally, simulation validation techniques were used to validate the model results. The outcomes of the simulation model were consistent with the actual observed phenomenon and the data gathered with the help of the subject matter experts.

6.2 Overview of ESIM

This thesis described the development of the Education System Intervention Modeling framework (ESIM), for modeling an intervention in a school system and identifying the key attributes affecting intervention sustainability. This was the first instance of such a modeling framework being developed to analyze interventions in the education system. The framework included agent-based modeling and social network analysis techniques, coupled with sensitivity analysis and factorial sampling using the Method of Morris. It also utilized the conceptual groundwork from education research via the three gaps: policy management, capability gap, and cultural gap; in order to define the acceptable zone and quantify the probability of sustainability of the intervention.

ESIM is composed of four different phases: *model definition*, *model design*, *model analysis*, and *model validation*. In the *model definition* phase, the problem being modeled is defined and the scale of analysis and the temporal scale at which the model is being built are identified. Then the agents, agents' attributes, and environment being modeled are determined. Finally, the criteria to define the sustainability of implementation of the particular intervention being modeled is determined, which populates the acceptable end states of the system, along with a method to quantify the probability of sustainability/risk of implementing the intervention. In the *model design* phase, first the agents' behaviors being modeled are listed. Then the agent-based model rules which govern the changes in agents' attributes and environment are developed. In the framework developed, concepts from discrete event simulation, Markov chains, and social network analysis are used to design the agent-based model rules. Finally, data collection and implementation of the computer simulation model are discussed in this phase. In the *model analysis* phase, first the simulation results are reviewed and sensitivity analysis is conducted with respect to the different input variables and parameters. Then the Method of Morris is applied to identify the inputs which have the largest impact on the output of interest. In the *model* *validation* phase, steps are carried out to verify and validate the model built using this framework. Both conceptual model validation and computer simulation model verification steps are discussed here. Collaboration between industrial and systems engineers, educational researchers, practitioners, and public policy researchers is important to build a model which is useful while analyzing the barriers and enablers of the particular intervention being modeled. Barriers are those attributes which hinder the implementation of the intervention and enablers are those attributes which facilitate the implementation of the intervention. Insights gained out of applying the ESIM framework to the SLIDER case study are discussed next.

6.3 SLIDER Case Study

ESIM was applied to model the *Science Learning: Integrating Design, Engineering, and Robotics* (SLIDER) intervention which was implemented across three schools in three different school districts. This was a five year intervention with varying outcomes across the three schools. School 3, which was in a suburban school district, was treated as a control variable since the system state of School 3 was already at the highest level. The system state of School 1 and School 2 was compared to the state of School 3 during the implementation of the intervention in order to assess the gaps across the three dimensions: policy management, capability, and cultural. During the course of the intervention, School 1 was moving towards an acceptable state where the intervention could be sustained after support from Georgia Tech is removed. Whereas, in School 2, the system state was moving away from the acceptable zone. This is also signified through the change in the probability of sustainability of the intervention measure at the two schools. This led to an interesting comparison of the implementation of the intervention across the two schools. One of the reasons why the intervention did not perform as well in School 2 in comparison to School 1 was that there were external trends at School 2, beyond the control of the intervention agency Georgia Tech, which were making it difficult for Georgia Tech to successfully implement the intervention. Also, the support received from the administration at School 2 was low, which translated to low support for intervention across the school.

Sensitivity analysis shows that the parameters in the change equations used in the model are fairly robust. This is important because estimating the parameters with high levels of accuracy can be tricky, the parameters can have inherent variability, and this noise should not cause significant changes in outcomes. Overall, for successful implementation of this intervention, the attributes that were determined to be most important through the Method of Morris analysis are: teachers' inquiry teaching skill, content knowledge and self-efficacy, budget for professional development, relationship between administration and Georgia Tech, number of students in a class, class duration, space, supplies, and teacher turnover rate. Here all the attributes except the teacher turnover rate are enablers, and the teacher turnover rate is a barrier.

After modeling the intervention at each school separately, the schools were modeled together as well. Variable proportion of resource allocation from Georgia Tech towards the three schools was analyzed as a function of the three gaps that existed at each school during the course of the intervention. The micro-behavior within the schools, which affected the three gaps, helped in describing the macro-behavior of the system in terms of variable proportions of resource allocation decisions being made by Georgia Tech.

6.4 Key Insights

One of the key factors identified in this work for sustainability of an intervention in a given system is the turnover rate amongst the agents involved in the intervention, especially the agents which are critical to the implementation of the intervention. The teacher turnover rate can be linked to the sustainability of the intervention in both the EWB and the SLIDER case studies. In the EWB case study, when the lead teacher and the principal changed in the second year, the intervention was not successful. In the SLIDER case study as well, teacher turnover rate came out to be an important attribute via the Method of Morris analysis, affecting the sustainability of the intervention.

Another insight gained is that different interventions can have different factors that lead to its sustainability, and generalizing such factors for all interventions is one of the main causes for so many interventions in the education system not being effective. For the SLIDER case study, a subset of attributes was identified which played a key role in its sustainability. However, this subset of attributes is quite different from the set of attributes that were identified as playing a key role in the successful implementation of the EWB intervention. This confirms our initial motivation in building a framework that can be adapted to different settings - there is a need to build different models to analyze different types of interventions.

Finally, a major insight gained out of this work is that using industrial engineering, systems engineering, and operations research techniques in the educational domain can be very useful. There is a lack of rigorous models developed for the educational system. As presented in this work, the use of modeling techniques, to analyze the implementation of the interventions in the K-12 school system, can guide the resource allocation to those components of the system which are important for the sustainability of the intervention. This helps in better utilization of the money being invested in the K-12 educational system. Through the application of the ESIM framework, the probability of sustainability or the risk of implementation can also be quantified. Previously, apart from subject matter experts' knowledge, there did not exist a method to quantify the risk of implementation of an intervention in a given school system.

6.5 Policy Implications

It is a fact that any intervention being implemented in any given system, and not just the education system, has to be carried out with a limited set of resources. These resources can be a combination of time, money, skilled professionals, information etc. There is a dearth of policy models, especially in the educations systems intervention arena, which analyze the resource allocation problem while carrying out such interventions. The ESIM framework helps in addressing this gap by providing insights into the drivers and inhibitors of the implementation of these interventions. Key attributes of the agents, and their relationships are identified. Educational researchers can use these insights to allocate resources in a more informed manner to maximize the chances of the intervention being sustainable.

Quantitative models do not exist which analyze change in the school system during the implementation of interventions. As discussed in Chapter 2, there is a new research methodology in educational research called Design-Based Implementation Research (DBIR), which tries to answer the questions such as: What type of interventions work, in which school settings? How can an intervention be made more sustainable? What capacities should the school system improve to facilitate the intervention? Currently, educational and public policy researchers are trying to address these questions, but they are not experts in the modeling and analysis techniques which industrial and systems engineering researchers use. Industrial and systems engineering researchers have demonstrated the benefits of applying these modeling techniques across various domains, such as healthcare, humanitarian aid, supply chain, financial systems etc. The ESIM framework presented in this thesis is the first application of using industrial and systems engineering techniques to analyze interventions in the education system. The questions posed above by the DBIR community can be better answered if the education and public policy researchers work together with the industrial and systems engineering researchers, utilizing their knowledge about the modeling techniques. This is what is proposed by the ESIM framework.

But, in order to better apply modeling techniques such as agent-based modeling, and social network analysis (which have been used in ESIM), knowledge about the school context, agents' attributes, relationships, and how these change is required. Therefore, it is not possible to apply the industrial and systems engineering techniques in isolation from the educational and public policy research. Educational research provides knowledge about which factors affect the change in the agents' attributes and their relationships. In some cases, it is possible that there does not exist previous studies about a particular attribute of an agent. If that attribute comes out be an important attribute for an intervention's sustainability, then that would provide incentives to the educational research community to conduct more studies around that attribute. The public policy researchers provide knowledge about the sustainability of an intervention, which is very important in characterizing the acceptable final states of the school system at the end of the intervention. Also, it is important to collaborate with the practitioners of the school intervention, so that knowledge about the school context can be obtained. Therefore, the application of industrial and systems engineering techniques, to analyze interventions in the education system, demands partnerships to be developed between researchers from these different disciplines, as well as the practitioners of school interventions.

Another implication for educational and public policy researchers, involved in the implementation of interventions, is the characterization of the risk of intervening in a given school system. With the help of the probability measures $P_H(I)$ and $P(I)$, discussed in *Section 4.2* of Chapter 4, the system can be categorized as either a *high-ris*k or a *low-risk* environment to intervene in. The changes taking place, during the intervention, in the three gaps: policy management, capability, and cultural; as well as the above probability measures, provide an assessment of the change in the system. The people carrying out the intervention can look at these changes to better understand the impact of the intervention on the particular school system.

Finally, although this falls under the category of future work, the framework developed is well situated to be used as a policy tool by intervention implementation agencies. As this framework is embedded in the application cycle of future interventions, a tool or an application can be developed in which one can adjust various inputs and parameters and then look at the change in certain outputs such as the three gaps and the probability of sustainability. Such a tool can be used throughout the course of the intervention to guide the decision makers regarding the impact of their decisions on the changes taking place in the school system. The benefit of developing this tool is that this would separate the educational and public policy researchers from all the industrial engineering, systems engineering, and operations research techniques being used.

6.6 Future Avenues of Research

One of the future avenues of research is to extend the framework to have a feature where an agent can choose the optimal mix of activities it wants to invest resources in. For example in the case of SLIDER intervention, Georgia Tech could be an agent who wants to decide which activities to carry out, and this can be done by using an optimization model. An initial approach to build the optimization model is presented below.

Parameters:

A	Set of activities
S_i	Set of attributes affected by activity i
u_j	Utility of attribute j
m_i	Money required to carry out activity i
t_i	Time required to carry out activity i
M	Total money available to spend
T	Total time available to spend

ME_j Mean effect of attribute j in the Method of Morris experiment

l_j Current level of attribute j

Decision variables:

$$x_i = \begin{cases} 1; activity\ i\ is\ carried\ out \\ 0; otherwise \end{cases}$$

Optimization Model:

$$max \sum_i \sum_{j \in S_i} u_j \cdot x_i$$

$$s.t.$$

$$\sum_i m_i \cdot x_i \le M; \qquad \text{(a)}$$

$$\sum_i t_i \cdot x_i \le T; \qquad \text{(b)}$$

$$u_j = f\left(ME_j, l_j\right) \qquad \forall j \qquad \text{(c)}$$

$$x_i = \{0,1\}$$

The objective function above maximizes the total utility obtained from carrying out the activities by the agent. Constraints (a) and (b) ensure that the total money and time spent on the activities is within the available budget and time constraint. Constraint (c) represents the utility function of attribute j. This function is not explicitly defined at the moment, it is proposed that the utility be represented as a function of the mean effect of attribute j, and the current level of attribute j in the system. Defining this utility function is another future research question.

The framework can be used to model more large scale interventions, at the district/state level, in the K-12 education system. The framework can also be used to model interventions that have a prominent teacher network. Social network analysis techniques, homophily and structural balance, demonstrated through the EWB case study can also be applied to model the teacher network. In addition, other concepts from social network analysis such as perception of expertise, reform activities, and proximity can be utilized while modeling teacher networks. The ESIM framework can also be extended to model interventions in the university education system. A lot of the characteristics of the university education system are similar to the K-12 education system. However, there are some changes/additions that would have to be incorporated in the modeling features in the framework in order to be able to model the university education system.

As the ESIM framework is applied to more case studies in the future, certain modeling features in the framework can be further tested. An important avenue of future research is to test the framework using different change equations and analyze the effect on model results. The change equations used in this framework were developed through a combination of the studies conducted in educational research and the modeling techniques such as agent-based modeling, social network analysis, and discrete-time Markov chains. These change equations can take different functional forms based upon the modeling techniques being used and assumptions being made. Applying the framework to various case studies, and using different types of change equations to model change in the system is a ripe avenue of future research. Another modeling feature to test can be the risk function to characterize the probability of sustainability of the intervention. In the ESIM framework, a modified logit model was proposed to calculate this risk measure. However, the functional form of this risk measure can be experimented with, in order to determine which function works the best, and under what conditions. Also, turnover in the agents is modeled as a random process. In the future, directed turnover in the agents can also be modeled (for example, intentionally replacing a low performing teacher in the school with a high performing one), if warranted by the case study. Finally, three gaps have been used to characterize sustainability of the intervention given the school context. These are policy management, cultural, and capability gaps. At the end of the ESIM application to the SLIDER case study, in Chapter 5, these gaps were also used to analyze the resource allocation decisions made by Georgia Tech while allocating additional resources across the three schools. In order to further develop confidence in the use of these gaps to make resource allocation decisions, such an analysis can be conducted across other interventions as well to which ESIM is applied.

Furthermore, to aid in the decision making process of future interventions, ESIM can also be used as a prospective tool at the beginning of the implementation of the intervention. Work on this has already started, ESIM has been incorporated into an NSF grant (EarSketch) [73]; an intervention that will be implemented across 30 different schools and will reach about 50 teachers and more than 1000 students over the 4 year grant period.

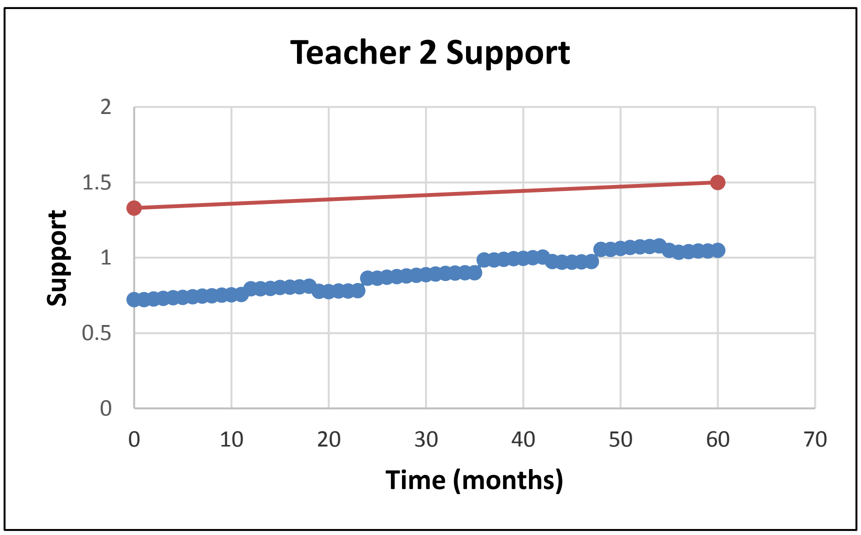
Through the application of ESIM, the risk of failure of an intervention can be reduced by improved allocation of resources towards those components of the system which play key roles in the success of the intervention. Increasing the success of interventions in the school system improves the educational outcomes in the school and increases the benefits gained from the investments being made in such interventions. It is hoped that the framework presented in this thesis becomes a starting point for the application of industrial and systems engineering in the educational domain, and encourages future research to be conducted at the intersection of these two fields.

APPENDIX A

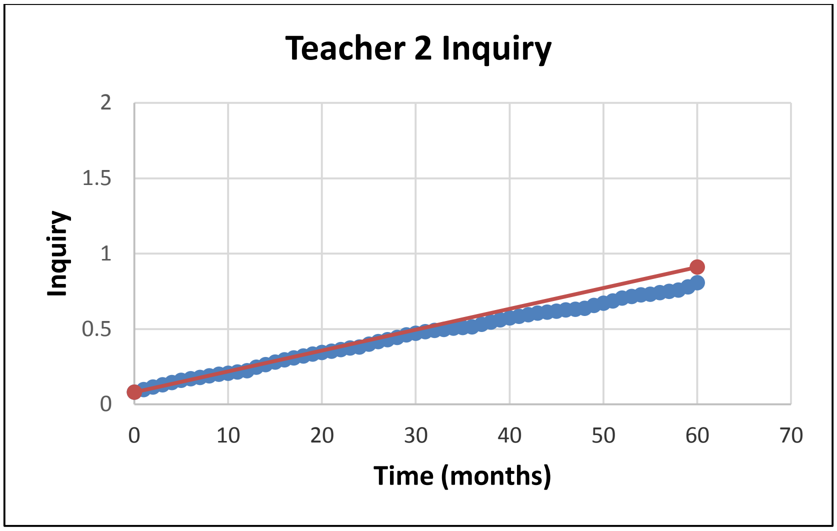
SIMULATION RESULTS: SLIDER

Additional simulation results, which were not presented in Chapter 5 because of brevity, are presented here. In all the graphs shown below, the blue line represents the model results over time, and the red line represents the actual measured values for these attributes. The red curve ties to the model validation phase of the framework, where the simulation results are validated using data.

School 1

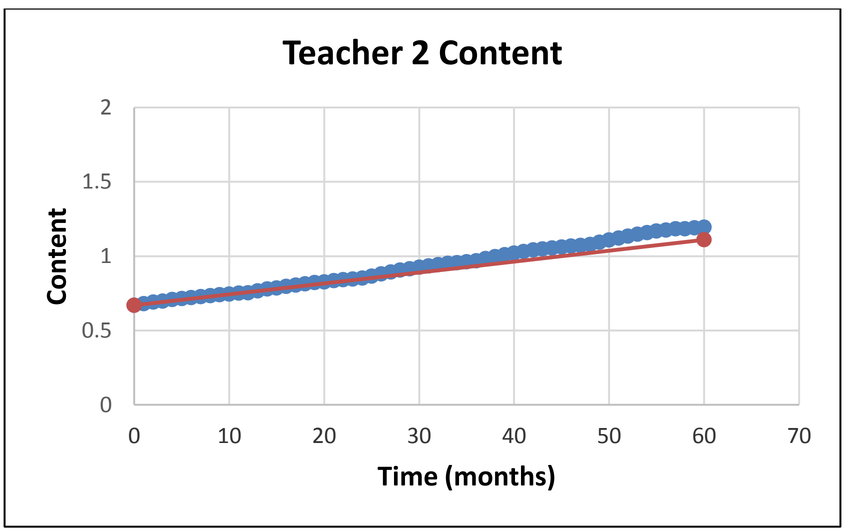


(a)



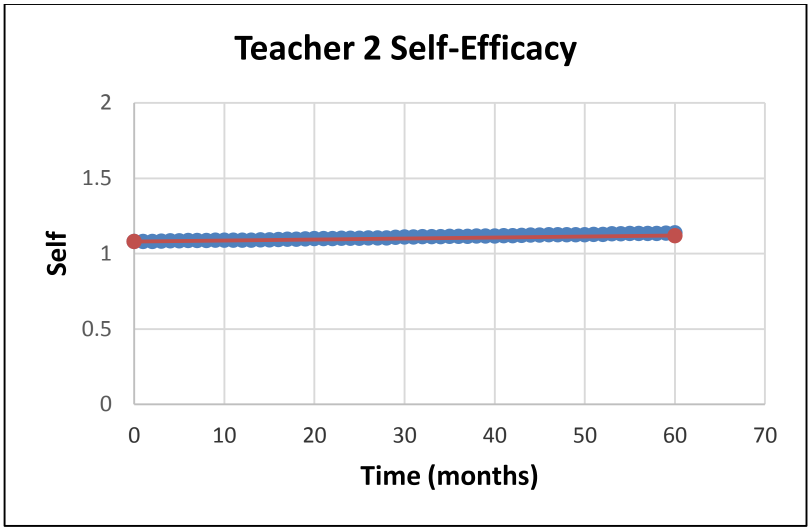
Teacher 2 Inquiry
Inquiry
2
1.5
1
0.5
0
0
10
20
30
40
50
60
70
Time (months)

(b)



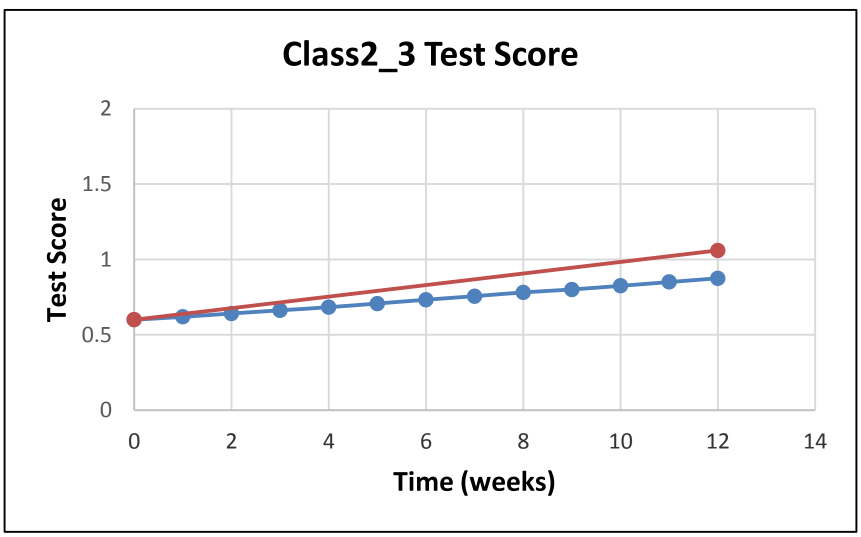
Teacher 2 Content
Content
2
1.5
1
0.5
0
0
10
20
30
40
50
60
70
Time (months)

(c)

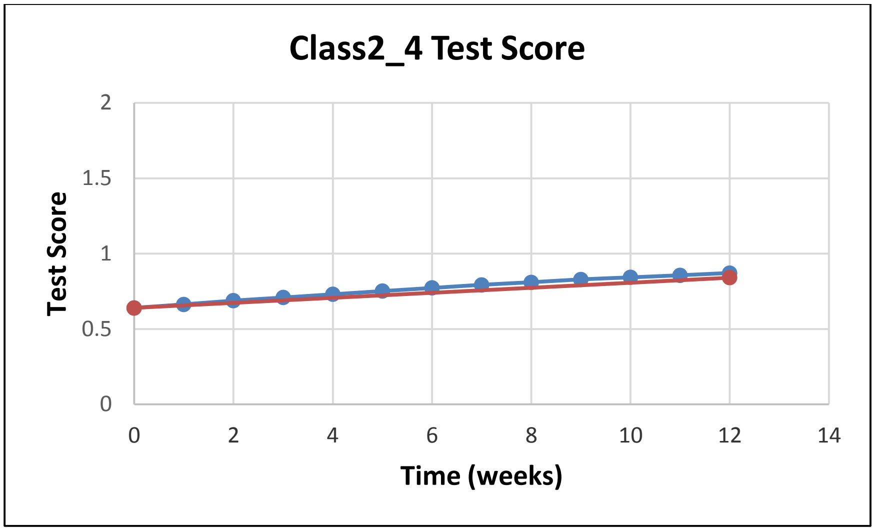


(d)

Figure 28: (a) Change in Teacher 2's support for intervention; (b) Change in Teacher 2's inquiry teaching skill; (c) Change in Teacher 2's content knowledge; (d) Change in Teacher 2's self-efficacy



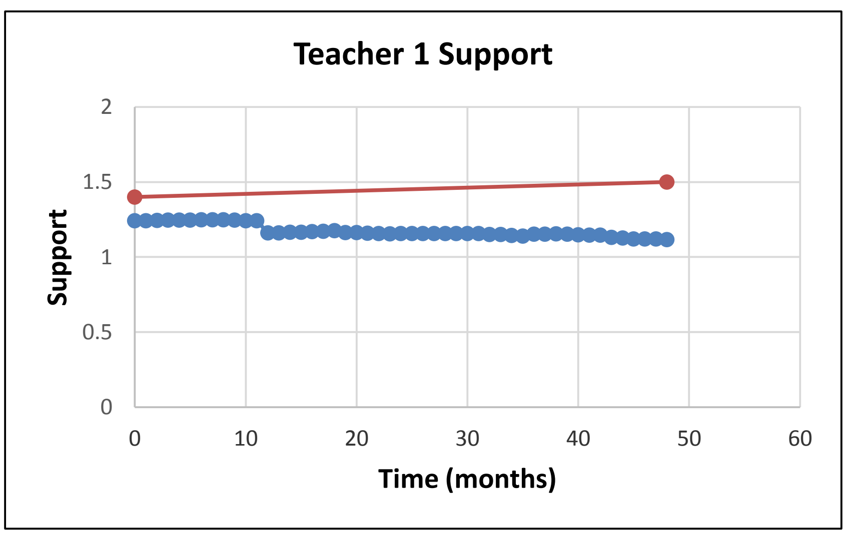
(a)



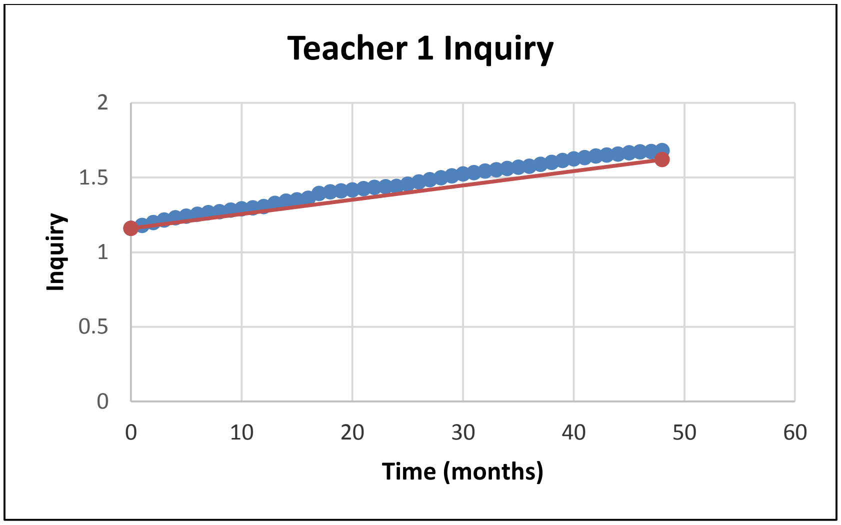
(b)

Figure 29: (a) Change in Class 2's SLIDER test scores in year 3; (b) Change in Class 2's SLIDER test scores in year 4

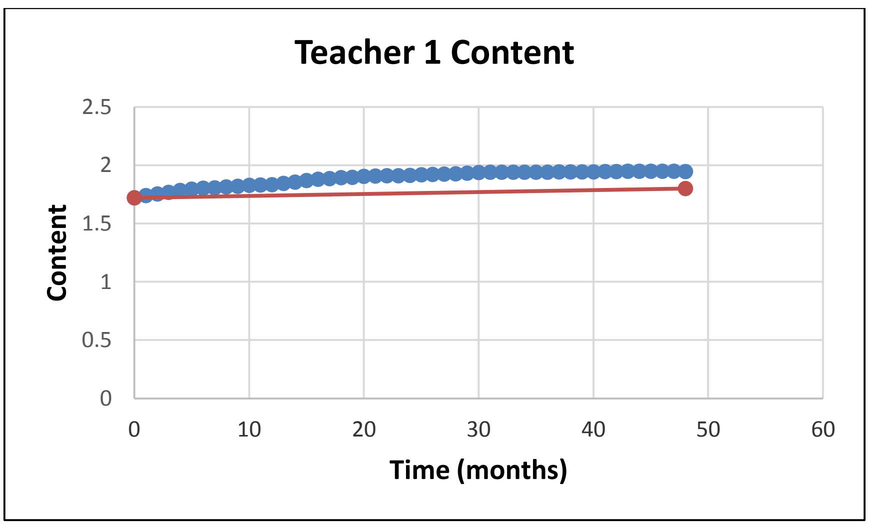
<u>School 2</u>



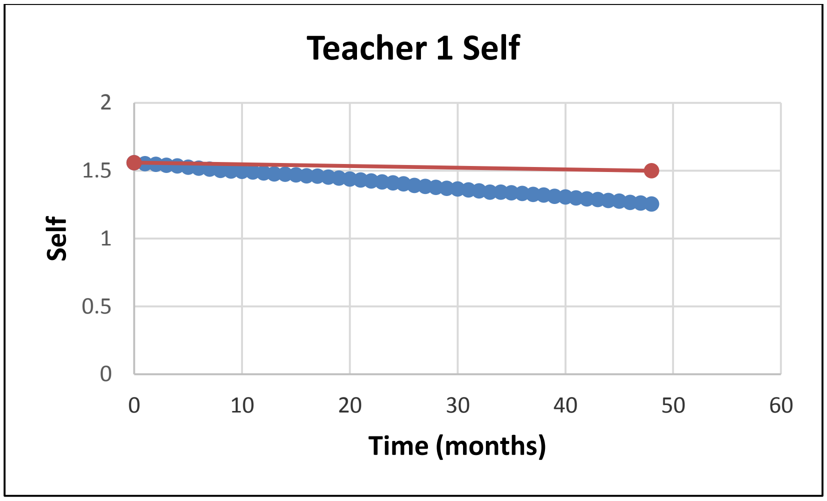
(a)


Teacher 1 Inquiry
Inquiry
2
1.5
1
0.5
0
0
10
20
30
40
50
60
Time (months)

(b)

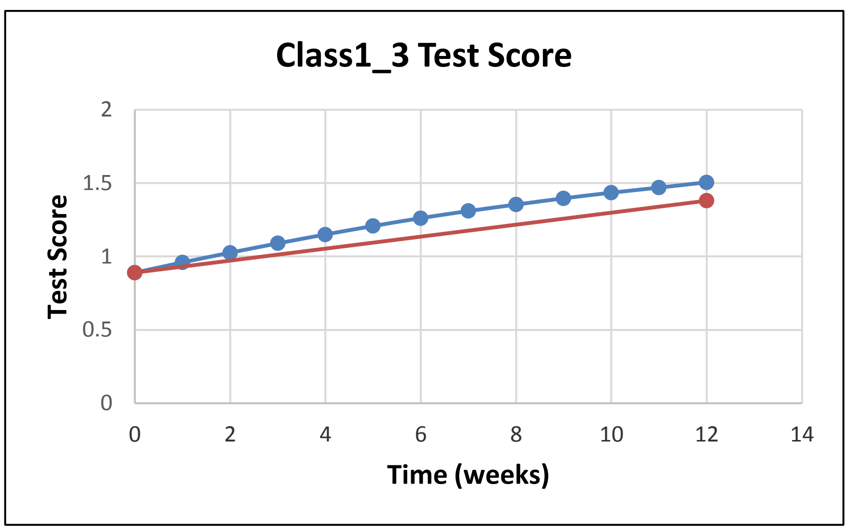

Teacher 1 Content
Content
2.5
2
1.5
1
0.5
0
0
10
20
30
40
50
60
Time (months)

(c)

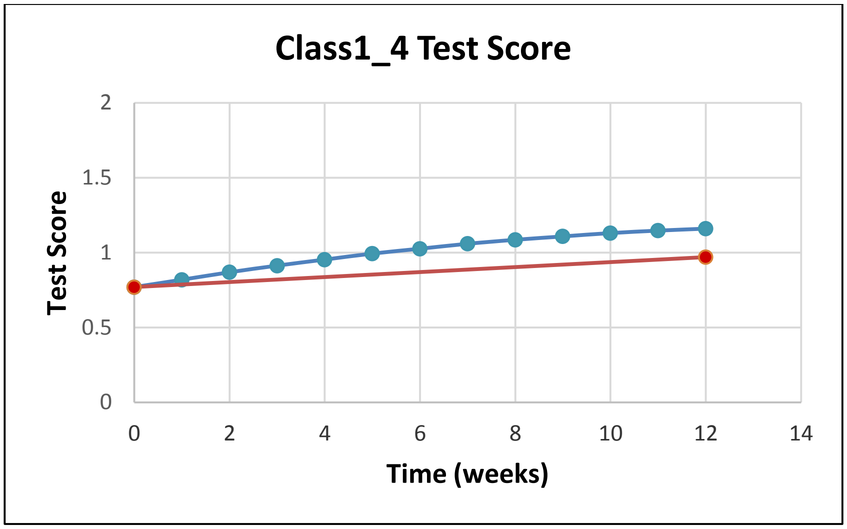


(d)

Figure 30: (a) Change in Teacher 1’s support for intervention; (b) Change in Teacher 1’s inquiry teaching skill; (c) Change in Teacher 1’s content knowledge; (d) Change in Teacher 1’s self-efficacy

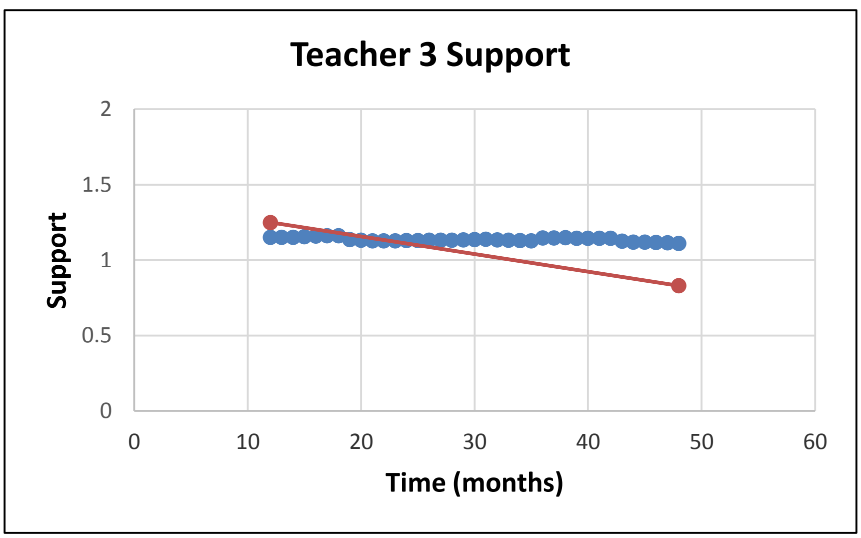


(a)

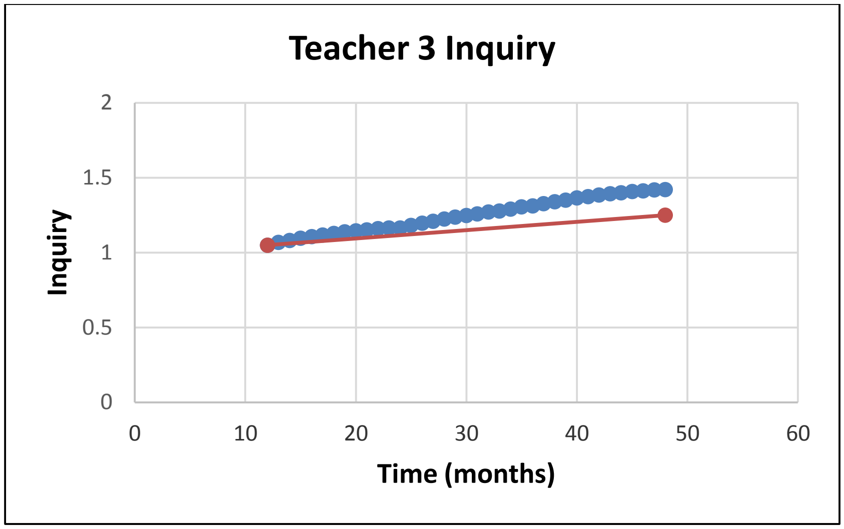


(b)

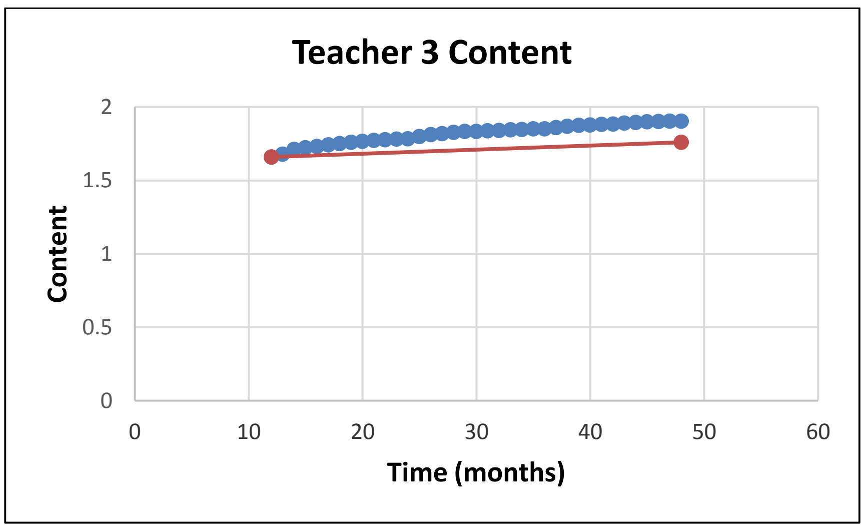
Figure 31: (a) Change in Class 1's SLIDER test scores in year 3; (b) Change in Class 1's SLIDER test scores in year 4



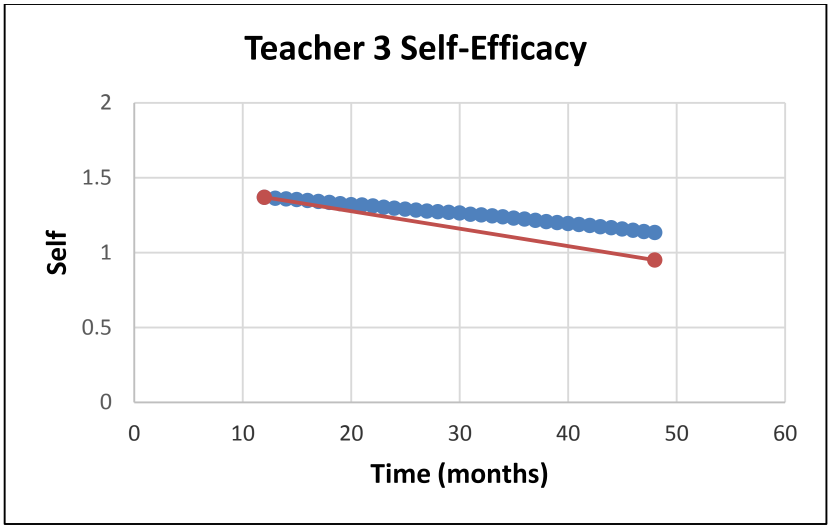
(a)


Teacher 3 Inquiry
Inquiry
2
1.5
1
0.5
0
0
10
20
30
40
50
60
Time (months)

(b)

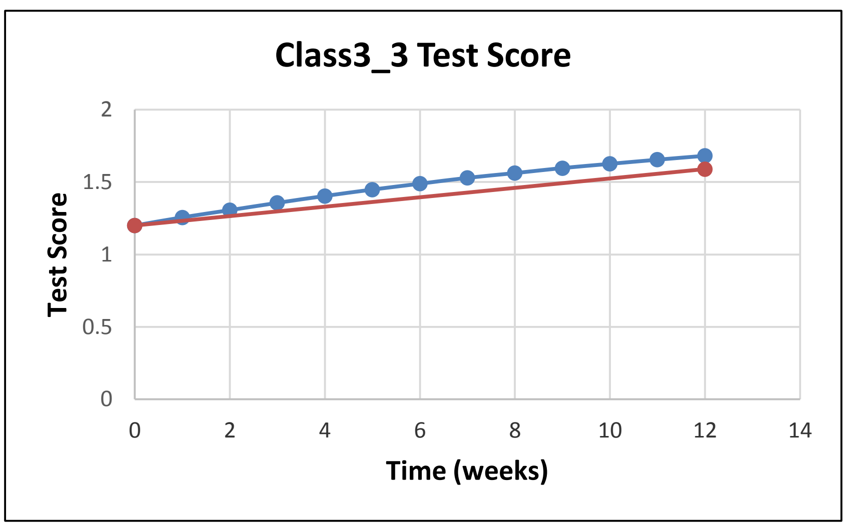

Teacher 3 Content
Content
2
1.5
1
0.5
0
0
10
20
30
40
50
60
Time (months)

(c)

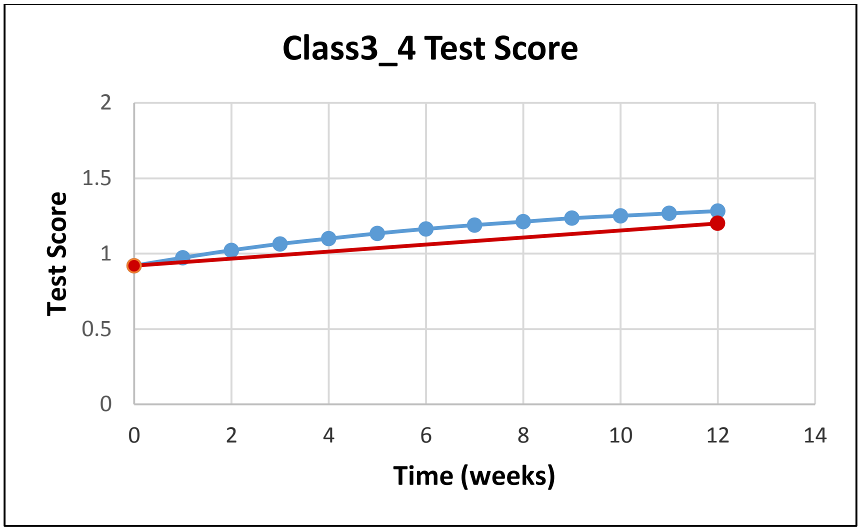


(d)

Figure 32: (a) Change in Teacher 3's support for intervention; (b) Change in Teacher 3's inquiry teaching skill; (c) Change in Teacher 3's content knowledge; (d) Change in Teacher 3's self-efficacy



(a)



(b)

Figure 33: (a) Change in Class 3's SLIDER test scores in year 3; (b) Change in Class 3's SLIDER test scores in year 4

REFERENCES

1. *No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat 1425 (2002)*.
2. Mital, P., R. Moore, and D. Llewellyn, *Analyzing K-12 education as a complex system.* Procedia Computer Science, 2014. **28**: p. 370-379.
3. Yam, B.Y., *Making things work: solving complex problems in a complex world.* 2005.
4. Sterman, J.D., *System dynamics modeling.* California management review, 2001. **43**(4): p. 8-25.
5. McLaughlin, M. and J. Talbert, *Reforming districts: How districts support school reform.* 2003.
6. Spillane, J.P., B.J. Reiser, and T. Reimer, *Policy implementation and cognition: Reframing and refocusing implementation research.* Review of educational research, 2002. **72**(3): p. 387-431.
7. Borman, K.M., et al., *Challenges for the future of comprehensive school reform.* Putting the pieces together: Lessons from comprehensive school reform research, 2004: p. 109-150.
8. Crawford, S.E. and E. Ostrom, *A grammar of institutions.* American Political Science Review, 1995: p. 582-600.
9. Weaver-Hightower, M.B., *An ecology metaphor for educational policy analysis: A call to complexity.* Educational researcher, 2008. **37**(3): p. 153-167.
10. Solomon, M.M., *Algorithms for the vehicle routing and scheduling problems with time window constraints.* Operations research, 1987. **35**(2): p. 254-265.

11. Santoso, T., et al., *A stochastic programming approach for supply chain network design under uncertainty.* European Journal of Operational Research, 2005. **167**(1): p. 96-115.

12. May, R.M., S.A. Levin, and G. Sugihara, *Complex systems: Ecology for bankers.* Nature, 2008. **451**(7181): p. 893-895.

13. Homer, J.B. and G.B. Hirsch, *System dynamics modeling for public health: background and opportunities.* American journal of public health, 2006. **96**(3): p. 452-458.

14. Jahre, M., et al., *Humanitarian logistics in disaster relief operations.* International Journal of Physical Distribution & Logistics Management, 2007. **37**(2): p. 99-114.

15. Sarkis, J., *Models for compassionate operations.* International Journal of Production Economics, 2012.

16. Sanchez, H.A., B. Wells, and J.M. Attridge. *Using system dynamics to model student interest in science technology, engineering, and mathematics*. in *Proceedings of The 27th International Conference of the System Dynamics Society*. 2009.

17. Allen, T.T. and N. Davis. *A simple agent-based social impact theory model of student STEM selection*. in *Simulation Conference (WSC), Proceedings of the 2010 Winter*. 2010. IEEE.

18. Groff, J.S., *Dynamic systems modeling in educational system design & policy*. New Approaches in Educational Research, 2013. **2**(2): p. 72-81.

19. Pedamallu, C.S., et al., *Investigating academic performance of migrant students: A system dynamics perspective with an application to Turkey.* International Journal of Production Economics, 2012. **139**(2): p. 422-430.

20. Pedamallu, C.S., et al., *A system dynamics model for improving primary education enrollment in a developing country.* Organizacija, 2010. **43**(3): p. 90-101.

21. Bonabeau, E. *Agent-based modeling: Methods and techniques for simulating human systems*. Proceedings of the National Academy of Sciences of the United States of America 2002 [cited 99 Suppl 3]; 7280-7287].

22. Moolenaar, N.M. and A.J. Daly, *Social Networks in Education: Exploring the Social Side of the Reform Equation.* American Journal of Education, 2012. **119**(1): p. 1-6.

23. Coburn, C.E., et al., *Supporting Sustainability: Teachers' Advice Networks and Ambitious Instructional Reform.* American Journal of Education, 2012. **119**(1): p. 137-182.

24. Moolenaar, N.M., *A Social Network Perspective on Teacher Collaboration in Schools: Theory, Methodology, and Applications.* American Journal of Education, 2012. **119**(1): p. 7-39.

25. Daly, A.J., *Social Network Theory and Educational Change*. 2010: ERIC.

26. Maroulis, S., et al., *Complex systems view of educational policy research.* Science(Washington), 2010. **330**(6000): p. 38-39.

27. Fixsen, D.L., et al., *Implementation research: A synthesis of the literature.* 2005.

28. Flay, B.R., et al., *Standards of evidence: Criteria for efficacy, effectiveness and dissemination.* Prevention science, 2005. **6**(3): p. 151-175.

29. Billig, S.H., L. Sherry, and B. Havelock, *Challenge 98: sustaining the work of a regional technology integration initiative.* British Journal of Educational Technology, 2005. **36**(6): p. 987-1003.

30. Blumenfeld, P., et al., *Creating usable innovations in systemic reform: Scaling up technology-embedded project-based science in urban schools.* Educational Psychologist, 2000. **35**(3): p. 149-164.

31. Education, C.f.I.T.i. *Technology Implementation in Schools: Key Factors to Consider*. December 2, 2014]; Available from: http://www.cited.org/index.aspx?page_id=187.

32. Elmore, R.F., *Getting to scale with good educational practice.* Harvard educational review, 1996. **66**(1): p. 1-27.

33. Ely, D.P., *Conditions that facilitate the implementation of educational technology innovations.* Journal of research on computing in education, 1990. **23**(2): p. 298-305.

34. Penuel, W.R., et al., *Organizing research and development at the intersection of learning, implementation, and design.* Educational Researcher, 2011. **40**(7): p. 331-337.

35. Penuel, W.R. and B.J. Fishman, *Large‐scale science education intervention research we can use.* Journal of research in science teaching, 2012. **49**(3): p. 281-304.

36. Railsback, S.F. and V. Grimm, *Agent-based and individual-based modeling: a practical introduction*. 2011: Princeton University Press.

37. Ghorbani, A., et al., *Maia: a framework for developing agent-based social simulations.* Journal of Artificial Societies and Social Simulation, 2013. **16**(2): p. 9.

38. Ostrom, E., *Understanding institutional diversity*. 2009: Princeton University Press.

39. Grimm, V., et al., *A standard protocol for describing individual-based and agent-based models.* Ecological modelling, 2006. **198**(1): p. 115-126.

40. Parker, D., et al., *'Illustrating a New 'Conceptual Design Pattern'for Agent-Based Models and Land Use via Five Case Studies: The MR POTATOHEAD Framework,''.* Agent-based Modeling in Natural Resource Management, eds. AL Paredes and CH Iglesias, Valladolid, Spain, 2008.

41. Mostafavi, A., D. Abraham, and D. DeLaurentis, *Ex-Ante Policy Analysis in Civil Infrastructure Systems.* Journal of Computing in Civil Engineering, 2013.

42. *Engineers Without Borders USA*. October 31, 2013]; Available from: http://www.ewb-usa.org/.

43. Nott, W., et al., *Engineers Without Borders–USA, Learning Through Humanitarian Service to Underdeveloped Communities.*

44. Tschannen-Moran, M. *Survey Instruments to Help You in Your Investigations of Schools*. November 4, 2013]; Available from: http://wmpeople.wm.edu/site/page/mxtsch/researchtools.

45. Oliver, N.M., B. Rosario, and A.P. Pentland, *A Bayesian computer vision system for modeling human interactions.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2000. **22**(8): p. 831-843.

46. Group, T.C.S.M., *Modelling in Healthcare*. 2010: American Mathematical Society.

47. Morris, M.D., *Factorial sampling plans for preliminary computational experiments.* Technometrics, 1991. **33**(2): p. 161-174.

48. Pahl, G., et al., *Engineering design: a systematic approach*. Vol. 157. 2007: Springer Science & Business Media.

49. Bar-Yam, Y., *Dynamics of complex systems*. Vol. 213. 1997: Addison-Wesley Reading, MA.

50. Llewellyn, D.C., et al., *Analyzing K-12 Education as a Complex System*, in *120th ASEE Annual Conference & Exposition*. 2013: Atlanta, GA.

51. Llewellyn, D., *Inquire within*. 2013: SAGE Publications.

52. Rogers, E.M., et al., *Complex adaptive systems and the diffusion of innovations.* The Innovation Journal: The Public Sector Innovation Journal, 2005. **10**(3): p. 1-26.

53. Alam, M.T. and M.S. Farid, *Factors affecting teachers motivation.* International journal of Business and social science, 2011. **2**(1): p. 298-304.

54. Duckworth, A.L. and P.D. Quinn, *Development and validation of the Short Grit Scale (GRIT–S).* Journal of personality assessment, 2009. **91**(2): p. 166-174.

55. Gillespie, M. and J. Noble, *Factors affecting student persistence: A longitudinal study*. Vol. 92. 1992: American College Testing Program Iowa City, Iowa.

56. Newmann, F.M., R.A. Rutter, and M.S. Smith, *Organizational factors that affect school sense of efficacy, community, and expectations.* Sociology of Education, 1989: p. 221-238.

57. Sirin, S.R., *Socioeconomic status and academic achievement: A meta-analytic review of research.* Review of educational research, 2005. **75**(3): p. 417-453.

58. Zeyu, X.a.A.S., *Estimating Teacher Contributions to Student Learning: The Role of the School Component.* 2013.

59. Zimmerman, B.J., A. Bandura, and M. Martinez-Pons, *Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting.* American educational research journal, 1992. **29**(3): p. 663-676.

60. Supovitz, J.A. and H.M. Turner, *The effects of professional development on science teaching practices and classroom culture.* Journal of research in science teaching, 2000. **37**(9): p. 963-980.

61. Davis, J.A., *Structural balance, mechanical solidarity, and interpersonal relations.* American Journal of Sociology, 1963: p. 444-462.

62. Heider, F., *The psychology of interpersonal relations.* 1958.

63. Gunawan, R., et al., *Sensitivity analysis of discrete stochastic systems.* Biophysical Journal, 2005. **88**(4): p. 2530-2540.

64. Snyder, L.V., *Supply chain robustness and reliability: Models and algorithms*. 2003, Northwestern University.

65. Conigliaro, R.A., A.A. Kerzhner, and C.J. Paredis. *Model-Based Optimization of a Hydraulic Backhoe Using Multi-Attribute Utility Theory*. in *SAE World Congress*. 2009.

66. Pace, D.K., *Ideas about simulation conceptual model development.* Johns Hopkins APL Technical Digest, 2000. **21**(3): p. 327-336.

67. Sargent, R.G. *Validation and verification of simulation models*. in *Simulation Conference, 2004. Proceedings of the 2004 Winter*. 2004. IEEE.

68. Naylor, T.H. and J.M. Finger, *Verification of computer simulation models.* Management Science, 1967. **14**(2): p. B-92-B-101.

69. Marion Usselman, J.C.A., Donna C. Llewellyn, Michael Ryan, *Science Learning: Integrating Design, Engineering, and Robotics (SLIDER).* 2009(NSF Project Number: 0918618).

70. Staw, B.M., L.E. Sandelands, and J.E. Dutton, *Threat rigidity effects in organizational behavior: A multilevel analysis.* Administrative science quarterly, 1981: p. 501-524.

71. *School Digger*. February 11, 2015]; Available from: http://www.schooldigger.com/go/GA/schools/0204002034/school.aspx?t=tbStudents&st=tbSTR.

72. *Georgia Department of Education*. February 11, 2015]; Available from: http://www.gadoe.org/Pages/Home.aspx.

73. Freeman, J., B. Magerko, and D. Edwards, *EarSketch: An authentic, studio-based STEAM approach to high school computing education.* NSF Abstract Award # 1417835, 2014.